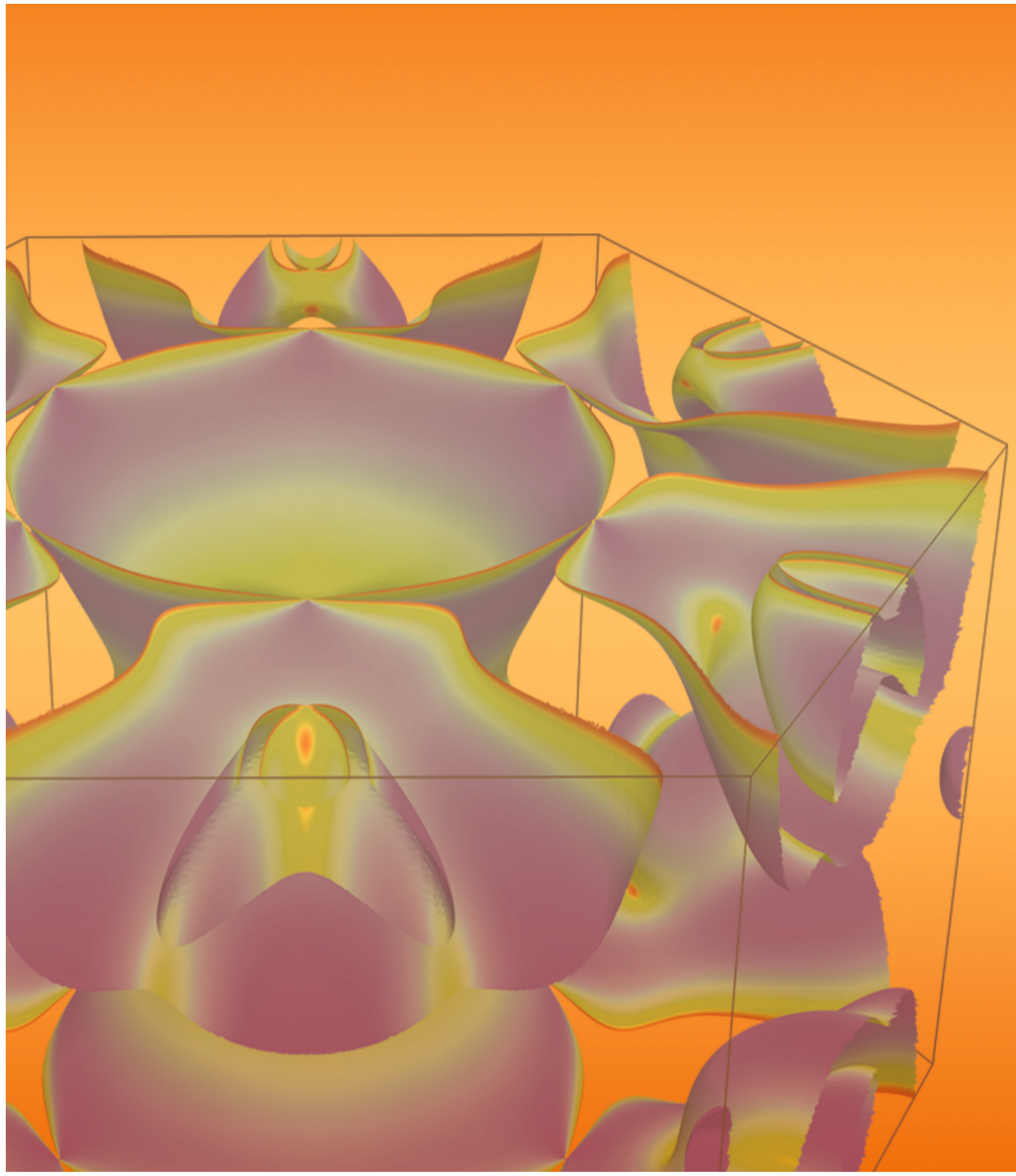
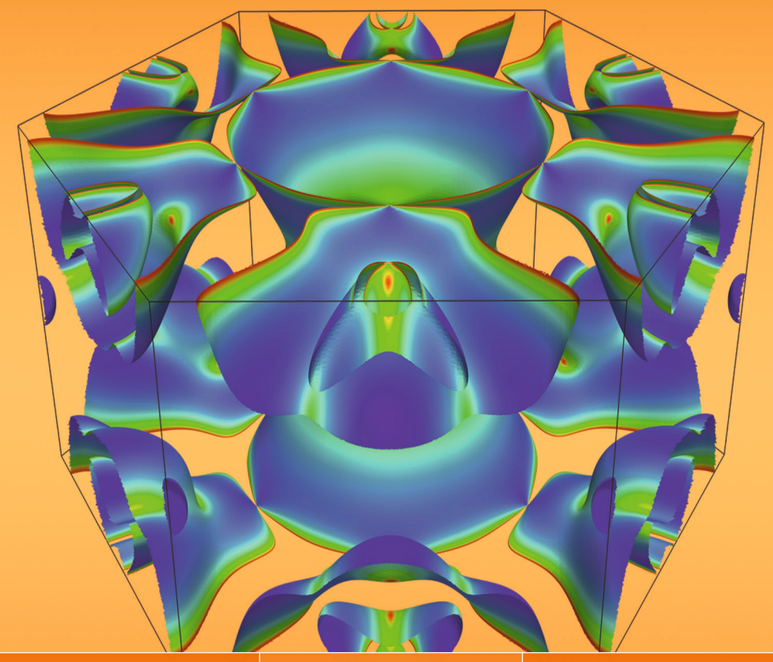


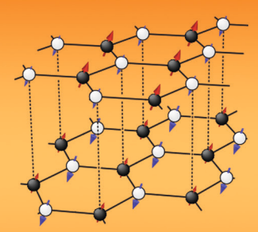
45th IFF Spring School 2014
Computing Solids
Models, ab-initio methods
and supercomputing

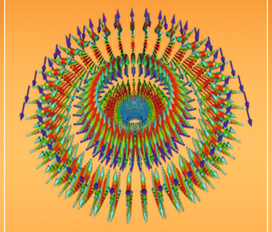
Lecture Notes

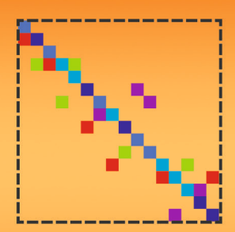
JÜLICH
FORSCHUNGSZENTRUM

Schlüsseltechnologien
Key Technologies









Schlüsseltechnologien / Key Technologies
Band / Volume 74
ISBN 978-3-89336-912-6

JÜLICH
FORSCHUNGSZENTRUM

74

Member of the Helmholtz Association

45th IFF Spring School 2014

Computing Solids

Models, ab-initio methods and supercomputing

Lecture Notes

JÜLICH
FORSCHUNGSZENTRUM

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies Band / Volume 74

Forschungszentrum Jülich GmbH
Institute for Advanced Simulation
Peter Grünberg Institute
Jülich Centre for Neutron Science
Institute of Complex Systems

Lecture Notes of the
45th IFF Spring School 2014

Stefan Blügel, Nicole Helbig, Volker Meden, Daniel Wortmann (Eds.)

Computing Solids
Models, ab-initio methods and supercomputing

The Spring School was organized
by the Institute for Advanced Simulation
and the Peter Grünberg Institute
of the Forschungszentrum Jülich
on March 10 – 21, 2014.

In collaboration with
universities, research institutes and industry.

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies Band / Volume 74

ISSN 1866-1807 ISBN 978-3-89336-912-6

Bibliographic information published by the Deutsche Nationalbibliothek.
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Publisher:	Forschungszentrum Jülich GmbH IAS, PGI, JCNS, ICS 52425 Jülich Phone +49 (0)2461 61-6048 · Fax +49 (0)2461 61-2410
Cover Design:	Grafische Medien, Forschungszentrum Jülich GmbH
Printer:	Schloemer + Partner GmbH, Düren
Copyright:	Forschungszentrum Jülich 2014
Distributor:	Forschungszentrum Jülich GmbH Zentralbibliothek, Verlag 52425 Jülich Phone +49 (0)2461 61-5368 · Fax +49 (0)2461 61-6103 e-mail: zb-publikation@fz-juelich.de Internet: http://www.fz-juelich.de

Schriften des Forschungszentrums Jülich
Reihe Schlüsseltechnologien / Key Technologies Band / Volume 74

ISSN 1866-1807
ISBN 978-3-89336-912-6

Part A – Quantum mechanics of solids

Part B – Quantum mechanical many body methods of solids

Preface

The computation of solids is challenged by the mutual interaction of its constituting elements, the myriad electrons and ions. The complex interplay produces a continuous stream of new and unexpected phenomena and forms of matter. The extreme range of length, time, energy and entropy scales established in the solid state give rise to a broad range of materials and associated properties. Some solids exhibit useful collective phenomena, such as ferroelectricity, magnetism, superconductivity, in others exotic states of matter such as the heavy fermion state are taken on. Varying external parameters such as the pressure or the doping it is even possible to switch between different ordered phases. Certain classes of solids show interesting metal to insulator transitions or display transversal, quantum and non-equilibrium transport processes, to mention a few of the ubiquitous emergent phenomena. New exotic phases or quantum states may occur for solids in low dimensions or at the nano- and mesoscopic scales. There are literally hundreds of thousands of solids with mostly unexplored properties. Every day, new solids or solid-state systems are synthesised or grown and novel properties are discovered. These solids find applications as present and emergent materials with specially-designed functionalities on which scientific advances in neighbouring disciplines such as metallurgy, materials science, nano-science, chemistry and biology as well as the geo-science rests on. Downstream applications can be found in information technology, green energy, transportation and health, all of enormous benefits to our society.

Even to physicists trained in the reductionistic view on nature it sometimes appears to be a miracle that the formation and stability of all solids and their wealth of properties are encoded in the statistical physics and quantum theory of the many electrons in the solid interacting via the Coulomb potential. It is the Schrödinger equation of many electrons which provides the fundamental theoretical concept for the understanding of the large variety of emerging quantum phenomena and processes that could be exploited in future technological devices. The exact analytical or numerical solution of such a Schrödinger equation for a solid is not in sight. Instead, since the formulation of the quantum mechanical many-body problem it remains a challenge to capture the properties of interacting electrons of complex atomic systems like e.g. a crystalline solid by approximate practical methods or effective models with reasonable computational effort.

In the past decades powerful theoretical concepts and reliable and predictive computational models have been developed that allow effective approximations. They aim at reducing complexity while retaining those ingredients necessary for a reliable description of the physical effects of the system. The underlying approximations made may be roughly divided into three different classes: realistic model Hamiltonians, that are solved in part with sophisticated and highly specialised analytical or numerical quantum many-body methods such as renormalization group based techniques or quantum Monte Carlo, wave function based methods and *ab initio* density functional approaches.

Computing solids refers to the application of these computational models to the study and prediction of the physical behaviour of solids. It represents an extension of theoretical physics that is based on mathematical models. The concept of computing solids can be used to predict new phenomena, to explore the validity of new concepts, to design new experiments in order to test these new concepts or simply to generate insight. It can be applied to complement and analyse experiments. It provides a powerful alternative to the techniques of experimental science when phenomena are difficult to observe or not observable with currently available techniques or when measurements are difficult, dangerous, expensive or simply impractical. It can be

understood as a virtual microscope, or a virtual physics or chemistry lab.

Today, the field of computing solids is on the threshold of a new era. The outstanding advances in the development of the physical and computational models go along with important progress in computational algorithms and methods, an impressive exponential growth of speed of computers, and permanent advances in the science and technology of data-intensive computing. A large number of configurations or parameter sets can be explored by high-throughput computing and can be deposited in data bases that will be analysed with data-mining and data-analysis tools opening the route to new materials and faster materials design with (pre-)defined properties. The increase of computer power now provides capabilities to address questions one could not even think of addressing 10 years ago. Much more accurate computational models, e.g. the treatment of (more) realistic model Hamiltonians, or advanced density-functionals will become affordable to a much wider class of solids. Due to the length and the time scales of the systems investigated including non-equilibrium setups and the complexity of the interactions, this field has benefited tremendously and will further benefit from the exponential growth of computer resources.

Simulations are carried out on a large diversity of computing platforms ranging from commodity PCs to high-performance supercomputers that operate in part with a peak performance of 10^{15} floating point operations (1 petaFLOP) per second using advanced architectures such a massively parallel computers or including graphical processor units (GPU). Adapting existing computer codes or developing new codes for these new infrastructures is an increasingly pressing and demanding topic in any computational based research. The development and application of above methods on supercomputers becomes a challenging science in itself. Since this is a very fast developing field, supercomputing yesterday is commodity computing today. High-end computing tomorrow enables investigations heretofore impossible, inspiring accelerating advances in computing solids of vast depth and breath with compelling opportunities.

The IFF Spring School 2014 provides a comprehensive introduction to modern concepts, theories and methods enabling the theoretical and computational description of many-electron systems. The hallmark of this year's school is in the concept of discussing density functional theory type *ab initio* approaches together with a variety of specialised numerical and analytical many-body techniques for strongly correlated electron systems emphasising the new opportunities provided by high-performance computers with advanced architectures. The school links these three aspects to the study of emergent properties of solids and non-equilibrium quantum transport. We point out their application for the grand challenges in information technology and energy materials. The link is facilitated for example by discussing *ab initio* Wannier functions and their role in constructing realistic model Hamiltonians for many-body quantum problems, the relation of topological concepts of the electronic structure and transverse transport properties, or by discussing ballistic transport in the single-particle regime in comparison to dissipative quantum mechanics in open systems coupled to a reservoir. The guiding principle for the selection of the subjects was their timeliness for the topical problems of students and young researchers. This course covers methods relevant for computing solids, a subject which is rarely treated in textbooks in this form, including a vast spectrum of concepts, theoretical and computational models, numerical tools, descriptions and techniques, but also topical materials and phenomena. To prepare students for research at supercomputers, their current and future architectures and their computing principles are introduced. While it is impossible to explore all subjects to full extent, due to the restricted time span available, it is the central aim of this school to provide a solid introduction, a basic understanding and an overview of the fascinating field of describing the quantum effects in extended, meso-scale and nano-scale solids from an

effective single particle and many-body picture. It seeks to rise the awareness of the ever growing potential of computing. Examples of topical applications, such as graphene, skyrmions, iron pnictides, analysis of scanning tunnelling microscopy experiments, or transversal transport properties like the Hall effects, allow the students to connect to their own research.
The IFF Spring School 2014 addresses these topics on a graduate student level. The subjects are divided into four linked parts, *Quantum Mechanics of Solids*, *Quantum Mechanical Many Body Methods of Solids*, *Simulating Materials Properties* and *Computing*, *Numerics, Computer Architecture*: More explicitly, the lectures cover

- Density Functional Theory and Methods
- Many-Body Perturbation Theory
- Model Hamiltonians
- Renormalization Group Techniques
- Berry Phase Physics and Wannier Functions
- Nonequilibrium Quantum-Transport and Open Systems
- Application to Graphene, Magnetic Skyrmions, Iron Pnictides
- Materials Informatics and Design
- Simulation Techniques
- Parallel Computing

The preparation of the 45th IFF Spring School benefited from the long track of the Forschungszentrum Jülich of advancing research in the field of computational materials physics and materials science. The IFF Spring Schools were first brought into being in 1970 by the "Institut für Festkörperforschung (IFF)". At a very early stage, scientists here recognised the importance of computer simulation as a source of knowledge in condensed matter physics. In 2011, the IFF was dissolved as part of a restructuring process and the Peter Grünberg Institute (PGI) emerged. Together with the Jülich Supercomputing Centre (JSC), the German Research School for Simulation Sciences (GRS), established under the umbrella of the Jülich-Aachen Research Alliance (JARA), and the recently founded Institute for Advanced Simulation (IAS), the Forschungszentrum Jülich is experienced in method development, large-scale simulations, algorithm and advanced architectures. This sets the stage of this spring school and guided in part the choice of lectures.
The spring school is organised together with the institutes Theory of Statistical Physics (TSP) and Theoretical Solid State Physics (TSSP) of the physics department of RWTH Aachen University as part of the Jülich-Aachen Research Alliance (JARA). They are an asset to the spring school enabling the concept of a concerted density functional and quantum-many body approach to the computation of solids, through their profound experience in the development of quantum many-body methods and their application in model systems of correlated electrons. Methods as well as numerical tools (e.g. quantum Monte Carlo) used to investigate the physics of correlated materials or correlation effects in mesoscopic systems including the rapidly developing field of non-equilibrium phenomena, or simulation techniques such as *ab initio* molecular dynamics found their way directly into lectures.
This school could not take place without the help and dedication of many colleagues. We are grateful to all contributors from the Institute for Advanced Simulation (IAS), the Peter Grünberg Institute (PGI), the Jülich Supercomputing Centre (JSC) and the German Research School for Simulation Sciences (GRS), as well as the colleagues from the theory institutes of the RWTH Aachen University. At the school, lectures will be given by these researchers from Aachen and Jülich:

Adinetz, Andrey V., Dr. (JSC)
Bihlmayer, Gustav, Dr. (IAS-1)
Costi, Theo A., Dr. (IAS-3)
Friedrich, Christoph, Dr. (PGI-1)
Honerkamp, Carsten, Prof. (TSSP)
Koch, Erik, Prof. (GRS)
Lounis, Samir, Dr. (PGI-1)
Mazzarello, Riccardo, Prof. (TSSP)
Mokrousov, Yuriy, Prof. (IAS-1)
Şaşıoğlu, Ersoy, Dr. (PGI-1)
Schnurpfeil, Alexander, Dr. (JSC)
Tsukamoto, Shigeru, Dr. (IAS-1)
Wessel, Stefan, Prof. (TSSP)
Zeller, Rudolf, Dr. (IAS-3)

Betzinger, Markus, Dr. (IAS-1)
Blügel, Stefan, Prof. (PGI-1/IAS-1)
Di Napoli, Edoardo, Dr. (JSC)
Helbig, Nicole, Dr. (PGI-1)
Jakobs, Severin G., Dr. (TSP)
Ležaić, Marjana, Prof. (PGI-1)
Mavropoulos, Phivos, Dr. (IAS-1)
Meden, Volker, Prof. (TSP)
Pleiter, Dirk, Prof. (JSC)
Schmidt, Manuel, Prof. (TSSP)
Schoeller, Herbert, Prof. (TSP)
Wegewijs, Maarten R., Prof. (PGI-2)
Wortmann, Daniel, Dr. (IAS-1)

We are especially glad that several distinguished colleagues from external universities and research laboratories have agreed to contribute to the program of the school:

Curioni, Alessandro, Dr.	Mathematical & Computational Sciences, IBM Research – Zurich, Switzerland,
Hoppe, Hans-Christian, Dipl. Ing.	ExaCluster Lab, Intel Cooperation, USA
Madsen, Georg K.H., Dr.	ICAMS, Ruhr University Bochum, Germany
Verstraete, Mattieu J., Prof.	Department of Physics, University of Liège, Belgium

We would like to express our sincere thanks to all colleagues for their time, effort and enthusiasm, which they have put into the preparation of their lectures and lecture notes. We thank particularly Prof. Dirk Pleiter for his guidance and advice in preparing the section Computing, Numerics, Computer Architecture of the IFF Spring School. We are very grateful to the board of directors of the Forschungszentrum Jülich for the continuous organisational and financial support, which we have received for the realisation of the IFF Spring School and the production of this book of lecture notes. Finally, our special thanks go to Mr. Michael Beißel for the general management, the organisation and his efforts in public outreach of the school.

Stefan Blügel, Nicole Helbig, Volker Meden, Daniel Wortmann January 2014

A 1 Describing many-electron systems [1]

D. Wortmann, N. Helbig, S. Blügel
Peter Grünberg Institut
Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

As the spring school gives an overview of state of the art approaches to compute properties of solids, many approaches discussed later in the school rest firmly on the well-explored fundamentals of solid-state theory as presented in an uncountable amount of textbooks, lecture note and university classes. While one could question the necessity to add more to this pile of available material, we believe that for completeness some of these fundamental aspects should also be covered here. This said, one of course realises that the sheer extend of the field of solid state theory makes even the creation of a simple overview a huge task well beyond the scope of this manuscript. Therefore, we will only present a very narrow view, relevant for the methods and ideas developed in further contributions.

Our focus will not be on the tools available from statistical physics, continuum mechanics, chemistry or even engineering to theoretically describe solids or their properties. This lecture only provides an introduction into the problem of the electronic structure of solids. To this aim, we will start by looking at the fundamental quantum mechanical description of mater in terms of the many-body Schrödinger equation. Before actually treating interacting electrons we introduce the Born-Oppenheimer approximation which is commonly employed in situations where the movement of the ions can be separated from the electronic degrees of freedom. We will spend some time discussing the properties and the description of many non-interacting electrons, especially those due to their fermionic nature in section 2. We then briefly review a few basic concepts and tools in a single-particle description of electronic systems in section 3. These include fundamental properties, such as symmetries, which can be exploited to simplify the calculations, and also key quantities that are commonly calculated like the bandstructure of solids or the density of states. In section 4 we return to the problem of interacting electrons and give a short overview of possible approximate methods commonly used to calculate electronic properties of solids and molecules.

As this chapter does not provide an in-depth introduction into any of the topics covered here, we refer for all the details to the large variety of textbooks on atomic and molecular physics, many-particle theory, solid state physics, or quantum chemistry, for example [1, 2, 3, 4]. The discussion of approximate methods for treating interacting systems in section 4 is intended to provide a first glance at the ideas discussed in much more detailed in other specialized lectures of the spring school and the corresponding manuscripts. As we will focus mainly on the single-particle picture, the contribution B1 of E. Koch can be seen as a complementary introduction more devoted to the many-body aspects of the field.

Throughout this chapter we employ atomic units where charge is measure in units of the charge of the electron, masses in units of the electron mass and energies in units of $\hbar$. In other words, we are setting $e = m = \hbar = 1$ in all formulas.

2 The Many-body Hamiltonian

The starting point of our discussion is the Hamiltonian of N identical electrons and N_n nuclei with masses M_I and charges Z_I, $1 \leq I \leq N_n$. The Hamiltonian then consists of three parts, one purely depending on electronic variables, one only containing nuclear coordinates and the third describing the interaction between electrons and nuclei,

$$\hat{H} = \hat{H}_e + \hat{H}_n + \frac{1}{2}\sum_{i=1}^{N}\sum_{I=1}^{N_n}\frac{Z_I}{|\mathbf{R}_I - \mathbf{r}_i|}. \tag{1}$$

The electronic Hamiltonian, $\hat{H}_e$, containing the kinetic energy of each electron $-\nabla^2/2$ and the electron-electron interaction between any two electrons at $\mathbf{r}_i$ and $\mathbf{r}_j$ given by $1/|\mathbf{r}_i - \mathbf{r}_j|$, reads

$$\hat{H}_e = \sum_{i=1}^{N} -\frac{\nabla_i^2}{2} + \frac{1}{2}\sum_{i\neq j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{2}$$

The nuclear Hamiltonian, $\hat{H}_n$, consists of the kinetic energy of each nucleus and the interaction between any two nuclei at positions $\mathbf{R}_I$ and $\mathbf{R}_J$, i.e.

$$\hat{H}_n = \sum_{I=1}^{N_n} -\frac{\nabla_I^2}{2M_I} + \frac{1}{2}\sum_{I\neq J}^{N_n} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{3}$$

In this non-relativistic formulation of the Hamiltonian is not spin-dependent as it would be if e.g. the spin-orbit interaction was included.

2.1 Born-Oppenheimer approximation

Using the Hamiltonian (1) the Schrödinger equation

$$\hat{H}\Psi_{total}(\{\mathbf{R}\}, \mathbf{r}_1..\mathbf{r}_N) = E\Psi_{total}(\{\mathbf{R}\}, \mathbf{r}_1..\mathbf{r}_N) \tag{4}$$

yields the combined wavefunction Ψ_{total} for the nuclei and the electrons with $\{\mathbf{R}\}$ denoting the complete set of ionic coordinates $\mathbf{R}_1...\mathbf{R}_{N_n}$. Due to the mass of the nuclei being much larger than the electron mass, $M_I \gg m$ one can effectively decouple the electronic and ionic coordinates. For fixed position of the ions the kinetic energy of the nuclei is zero. The ion-ion interaction and the electron-ion interaction depend parametrically on the position of all nuclei but the only remaining variables are the electronic coordinate $\mathbf{r}_1...\mathbf{r}_N$. Therefore, we can rewrite the two terms which contain ionic coordinates as an external potential $v_{\text{ext}}(\mathbf{r})$ and the remaining electronic Hamiltonian reads as

$$\hat{H} = \sum_{i=1}^{N} \left(-\frac{\nabla_i^2}{2} + v_{\text{ext}}(\mathbf{r}_i) \right) + \frac{1}{2}\sum_{i\neq j}^{N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \tag{5}$$

The corresponding Schrödinger equation $\hat{H}\Psi_k(\mathbf{r}_1...\mathbf{r}_N; \{\mathbf{R}\}) = E_k(\{\mathbf{R}\})\Psi_k(\mathbf{r}_1...\mathbf{r}_N; \{\mathbf{R}\})$ then yields the electronic wavefunctions Ψ_j and their energy eigenvalues E_j which both parametrically depend on the ionic positions $\{\mathbf{R}\}$. Since the set of electronic wavefunctions is complete, we can expand the total wavefunction Ψ_{total} as

$$\Psi_{total}(\{\mathbf{R}\}, \mathbf{r}_1..\mathbf{r}_N) = \sum_{k=1}^{\infty} F_k(\{\mathbf{R}\})\Psi_k(\mathbf{r}_1...\mathbf{r}_N; \{\mathbf{R}\}). \tag{6}$$

Plugging this expansion of the total wavefunction into the Schrödinger equation (4) we obtain

$$\sum_{k=1}^{\infty}\sum_{I=1}^{N_n} -\frac{\nabla_I^2}{2M_I} F_k(\{\mathbf{R}\})\Psi_k(\mathbf{r}_1...\mathbf{r}_N; \{\mathbf{R}\}) = \sum_{k=1}^{\infty}(E - E_k(\{\mathbf{R}\}))F_k(\{\mathbf{R}\})\Psi_k(\mathbf{r}_1...\mathbf{r}_N; \{\mathbf{R}\}). \tag{7}$$

We multiply this equation with $\Psi_l(\mathbf{r}_1...\mathbf{r}_N;\{\mathbf{R}\})$ and integrate over all electronic coordinates. Due to the orthonormality of the electronic wavefunctions for each fixed ionic position $\{\mathbf{R}\}$ this yields

$$\sum_{k=1}^{\infty}\sum_{I=1}^{N_n}\int d^3r_1...\int d^3r_N\Psi_l^*(\mathbf{r}_1...\mathbf{r}_N;\{\mathbf{R}\})\frac{-\nabla_I^2}{2M_I}F_k(\{\mathbf{R}\})\Psi_k(\mathbf{r}_1...\mathbf{r}_N;\{\mathbf{R}\}) = (E-E_l(\{\mathbf{R}\}))F_l(\{\mathbf{R}\}). \tag{8}$$

The Born-Oppenheimer approximation now consists in neglecting the derivatives of the electronic wave functions $\Psi_k(\mathbf{r}_1...\mathbf{r}_N;\{\mathbf{R}\})$ with respect to the ionic coordinates $\mathbf{R}_I$. In this case Eq. (8) decouples and yields the following equation for the ionic wavefunction F_l

$$\left[\sum_{I=1}^{N_n}-\frac{\nabla_I^2}{2M_I}+E_l(\mathbf{R}_1...\mathbf{R}_{N_n})-E\right]F_l(\mathbf{R}_1...\mathbf{R}_{N_n})=0, \tag{9}$$

where we have reverted to writing the complete set of ionic coordinates.
Neglecting the derivatives of the electronic wavefunction with respect to the ionic coordinates is justified in many cases because the electronic excitation energies are much larger than their ionic counterparts. The former lie in the visible or ultraviolet range while the latter are in the infrared. Therefore, a small change in the ionic coordinates around their equilibrium positions leaves the electrons mostly unaffected. However, there are situations in which a small change in ionic positions yields a drastic change in the electronic wavefunction, for example when two electronic states are nearly degenerate and one or the other has a lower energy upon a small change in ionic positions. Schematically this situation is plotted in Fig. 1 where the dependence of the two energy levels E_1 and E_2 is given as a function of the ionic positions $\{\mathbf{R}\}$ calculated using the Born-Oppenheimer approximation. Including the neglected terms in a perturbative treatment one can show that the crossing of the two levels is actually turned into an avoided crossing signified by the dashed lines in the plot.
In Eq. (9) $E_l(\mathbf{R}_1...\mathbf{R}_{N_n})$ provides the interaction potential in which the nuclei move. It includes the direct ion-ion interaction as well as the effect of the electrons on the ions. Many of such potential surfaces $E_l(\mathbf{R}_1...\mathbf{R}_{N_n})$ exist but usually only the $l=1$ surface corresponding to the electronic ground state is relevant. In this interpretation one would expect the Born-Oppenheimer approximation to be valid only if the different higher $E_{l>1}(\mathbf{R}_1...\mathbf{R}_{N_n})$ are well separated from the ground state $E_{l=1}(\mathbf{R}_1...\mathbf{R}_{N_n})$

2.2 Non-interacting electrons

In the following we will assume that the Born-Oppenheimer approximation is valid and continue with the Hamiltonian for the electronic degrees of freedom Eq. (5). This Hamiltonian can be split further into two very different contributions

$$\hat{H}=\hat{H}_0+\hat{V}_{ee} \tag{10}$$

with

$$\hat{H}_0=\sum_{i=1}^{N}\hat{h}(\mathbf{r}_i)=\sum_{i=1}^{N}-\frac{\nabla_i^2}{2}+v_{\text{ext}}(\mathbf{r}_i) \tag{11}$$

and

$$\hat{V}_{ee}=\frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}. \tag{12}$$

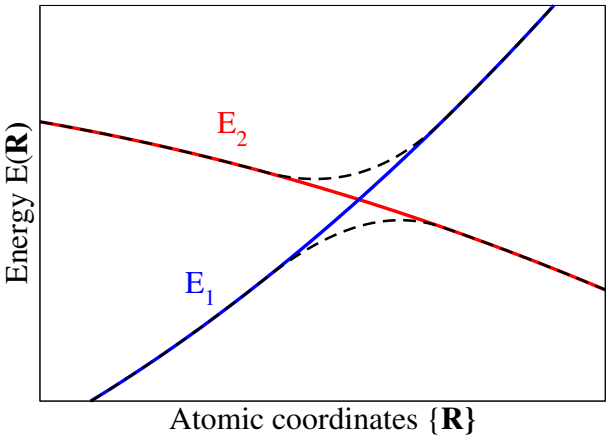


Fig. 1: *Schematic energy versus ionic coordinates plot. At the crossing of the two energy curves a small change in the ionic coordinates changes the electronic configuration dramatically causing the Born-Oppenheimer approximation to fail.*

$\hat{H}_0$ alone describes a set of N non-interacting electrons in the external potential, v_{ext}. Each of these non-interaction electrons is described by a single-particle Hamiltonian $\hat{h}$. The second term in Eq. (10), $\hat{V}_{ee}$, describes the interaction between the electrons. This term is the sum over all two-particle Coulomb interaction terms. As the Coulomb interaction is a strong interaction at short distances between two electrons and at the same time an interaction only slowly decaying with distance this term makes even an approximate solution of the stationary many-body Schrödinger equation

$$\left(\hat{H}_0 + \hat{V}_{ee}\right) \Psi(\mathbf{r}_1, ...\mathbf{r}_N) = E\Psi(\mathbf{r}_1, ...\mathbf{r}_N) \tag{13}$$

extremely difficult.
To gain some insight into the problem of interacting electrons one might first look at the situation of non-interacting or independent electrons, e.g. neglecting the electron-electron interaction ($\hat{V}_{ee} = 0$). In this case the total Hamiltonian is simply given by

$$\hat{H}_0 = \sum_{i=1}^{N} \hat{h}(\mathbf{r}_i) \tag{14}$$

and whence a solution of the stationary Schrödinger equation

$$\hat{H}_0 \Psi(\mathbf{r}_1, ...\mathbf{r}_N) = E\Psi(\mathbf{r}_1, ...\mathbf{r}_N) \tag{15}$$

is given by the simple product ansatz

$$\Psi(\mathbf{r}_1, ...\mathbf{r}_N) = \psi_1(\mathbf{r}_1) \dots \psi_N(\mathbf{r}_N) \tag{16}$$

where the ψ_i are eigenvectors of the single-particle Hamiltonian $\hat{h}$ with eigenvalue ϵ_i

$$\hat{h}\psi_i = \epsilon_i \psi_i. \tag{17}$$

Obviously, this ansatz leads to

$$E = \sum_{i=1}^{N} \epsilon_i. \tag{18}$$

One should note, that we have introduced an index i for the single-particle wavefunctions ψ_i which labels the different eigenfunctions of the single-particle Hamiltonian $\hat{h}$ in the construction of the many-body wavefunction in Eq. (16). Even though the particles are assumed to be non-interacting, we have to consider that we are dealing with a quantum mechanical problem in which the Pauli exclusion principle does not allow two electrons (fermions) to occupy the same state. Hence, the state of Eq. (16) should be constructed such that all the i are different. Obviously, the ground-state energy of the system will be given by combining the lowest N eigenstates ψ_i of $\hat{h}$ with eigenvalues $\epsilon_i \leq \epsilon_F$, where the highest energy of the single-particle states involved in the construction of the many-body wavefunction is called the Fermi-energy ϵ_F (cf. Sec. 3.2 for more details on the Fermi energy).
Because of the spin-$1/2$ character of the electrons the wavefunction must be assigned an additional spin-quantum number. This requires an extension of the notation we used so far to include a spin variable. Hence, we introduce $\mathbf{x} = (\mathbf{r}, \sigma)$ to denote the combination of the spatial and spin degrees of freedom. Additionally, in cases in which a spin-dependent single-particle potential appears it will be necessary to extend the index i which labels the different single-particle wavefunctions by a spin-variable.
Our ansatz for the many-body wavefunction of N non-interacting electrons still suffers from a serious problem in the treatment of the fermionic nature of the electrons. The naive fix of using every single-particle state only once does not fulfill the more general formulation of the exclusion principle which requires the many-body wavefunction to be anti-symmetric under the exchange of two particles

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_N) = -\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N). \tag{19}$$

This condition cannot be fulfilled by a simple product of single-particle states as used in Eq. (16). However, an anti-symmetric linear combination of such products can be constructed. This construction is known as a 'Slater determinant' as it can be expressed in terms of a determinant of a matrix containing the single-particle states

$$\begin{aligned} \Psi_{\text{Slater}}(\mathbf{x}_1 \ldots \mathbf{x}_N) &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \ldots & \psi_1(\mathbf{x}_N) \\ \vdots & \ddots & \vdots \\ \psi_N(\mathbf{x}_1) & \ldots & \psi_N(\mathbf{x}_N) \end{vmatrix} \\ &= \frac{1}{\sqrt{N!}} \sum_P (-1)^P P\left(\psi_1(\mathbf{x}_1) \ldots \psi_N(\mathbf{x}_N)\right). \end{aligned} \tag{20}$$

In this notation the sum is performed over all permutations P acting on the indices i of the ψ_i. The factor $(-1)^P$ ensures the required anti-symmetry. It can be shown that the Slater-determinants as given in Eq. (20) form an anti-symmetric solution of the non-interacting Schrödinger equation (15).

The spherical potential of the nucleus

While many difficulties of solving the many-electron Hamiltonian are linked to the two-particle electron-electron interaction term we should also shortly discuss the external potential. First one should recall that the main contribution to this term is the strong attractive Coulomb potential due to the nuclei. Hence, close to the position $\mathbf{R}_i$ of a nucleus of an atom the external potential will be

$$v_{\text{ext}}(\mathbf{r}) \approx \frac{Z}{|\mathbf{r} - \mathbf{R}_i|}$$

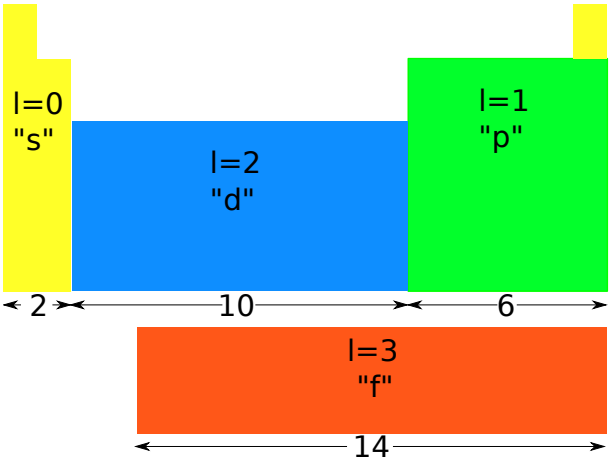


Fig. 2: *Schematics of the periodic table of the elements where only the blocks corresponding to different l-values are shown.*

and other terms like external fields will be completely negligible at small $r = |\mathbf{r} - \mathbf{R}_i|$. As this potential is spherically symmetric one can solve the corresponding one particle Schrödinger equation easily. This is discussed in basic quantum mechanic lectures and thus we will only shortly recap the results. The solution is written as product ansatz of a radial dependent function and a spherical harmonic

$$\psi(\mathbf{r}) = u_l(r) Y_{l,m}(\mathbf{r})$$

where $u_l(r)$ is a function of the distance from the nucleus only and the $Y_{l,m}$ depend only on the angular coordinates. This leads to a simple differential equation for the radial function

$$\left(-\frac{1}{2}\frac{d^2}{dr^2} - \frac{Z}{|\mathbf{r} - \mathbf{R}_i|} + \frac{l(l+1)}{2r^2} \right) r u_l(r) = \epsilon_l r u_l(r) \tag{21}$$

where the eigenvalues ϵ_l (and eigenfunctions) only depend on the $l-$value of the spherical harmonic. This has important consequences as for each l one has $2l + 1$ different values for m and therefore the eigenvalues are $2l + 1$ times degenerate. One usually uses the convention of calling $l = 0, 1, 2, 3$ solutions as s, p, d, f-states, respectively. It follows that there is one $s-$state, 3 $p-$states, 5 $d-$states and 7 $f-$states. Additionally, there are of course different solutions of Eq. (21) for a given l, which are counted by a so-called principle quantum number n. Together with the spin-variable σ, the four indices n, l, m, σ completely specify the solution of the radially symmetrical Schrödinger equation, where the energy eigenvalues are degenerate for the different m and σ. Even though we completely disregarded the electron-electron interaction in this discussion, this simple investigation already introduces many basic concepts of chemistry and can be used to motivate the ordering of the elements in the periodic table (c.f. Fig. 2). While the details of the ordering of the energy levels can not be understood from this discussion, the different blocks of elements with partly filled s, p, d or f states and the extend of these blocks can be understood. Concluding, one should stress, that the potential of the nuclei is one of the most relevant ingredients of the electronic Hamiltonian. In some sense, its effect is even stronger than those of the electron-electron interaction as the electronic charge density is smeared out. In other words, the electron-electron interaction mainly changes effects on a rather small energy scale while the high energy scale is set by the potential of the nuclei.

3 Key concepts in a single particle description of solids

The description of the electronic structure of a solid usually takes advantage of several basic concepts and properties of which we will now introduce a few. While some concepts are rather general, we will restrict our discussion to a description in terms of single particle states, i.e. we will assume that the fundamental description of the system is possible by considering wavefunctions ψ_i that are solutions of the Schrödinger equation

$$\hat{H}\psi_i(\mathbf{r}) = \left(-\frac{1}{2}\nabla^2 + V(\mathbf{r})\right)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}).$$

$V(\mathbf{r})$, the single particle potential will of course include the effects of the external potential, i.e. the effect of the nuclei. Additionally, some of the effects of the electron-electron interaction appearing in the full many-electron Schödinger equation (Eq. (5)) often are also incorporated into an "effective" or "mean-field" single particle potential. We will come back to this point in Sec. 4.3 in which we will discuss the reasons why the description in terms of single (=non-interacting) particles is so successful.

3.1 Symmetries and boundary conditions

Frequently, the use of symmetry allows for a significant simplification of the problem and the eigenstates and eigenvalues of the Hamiltonian can be classified according to their symmetry properties. Formally, in this situation one can define a symmetry operator $\hat{O}_M$ under which the Hamiltonian is invariant, i.e. the two operators $\hat{H}$ and $\hat{O}_M$ commute

$$\left[\hat{H}, \hat{O}_M\right] = \hat{H}\hat{O}_M - \hat{O}_M\hat{H} = 0. \tag{22}$$

Furthermore, it is possible to diagonalize $\hat{H}$ and $\hat{O}_M$ simultaneously. This means there exists a common set of eigenvectors such that

$$\hat{H}\psi_i = \epsilon_i\psi_i \tag{23}$$

and

$$\hat{O}_M\psi_i = m_i\psi_i. \tag{24}$$

One can then classify the common eigensolutions by their symmetry properties, i.e. by the eigenvalues m_i of the symmetry operator. This classification can be a powerful tool to make predictions about the eigenspectrum of the Hamiltonian. In particular the existence of symmetries often leads to degenerate eigenstates with the same energy eigenvalue ϵ_i. The investigation of the algebraic properties of the symmetry operations of the system, i.e. the group properties of the symmetries, hence is an important ingredient in quantum mechanics and can give very general properties of the system.

Simple example of a symmetry

As a most basic example we consider a situation in which the system exhibits a mirror plane such that $V(\mathbf{r}) = V(r_1, r_2, r_3) = V(-r_1, r_2, r_3)$ (c.f. Fig. 3). Thus, the Hamiltonian will not change if we replace $r_1 \leftrightarrow -r_1$ as the kinetic energy does not change by applying the mirror operation. Therefore, the symmetry operator $\hat{O}_M$ in this example simply replaces r_1 by $-r_1$.

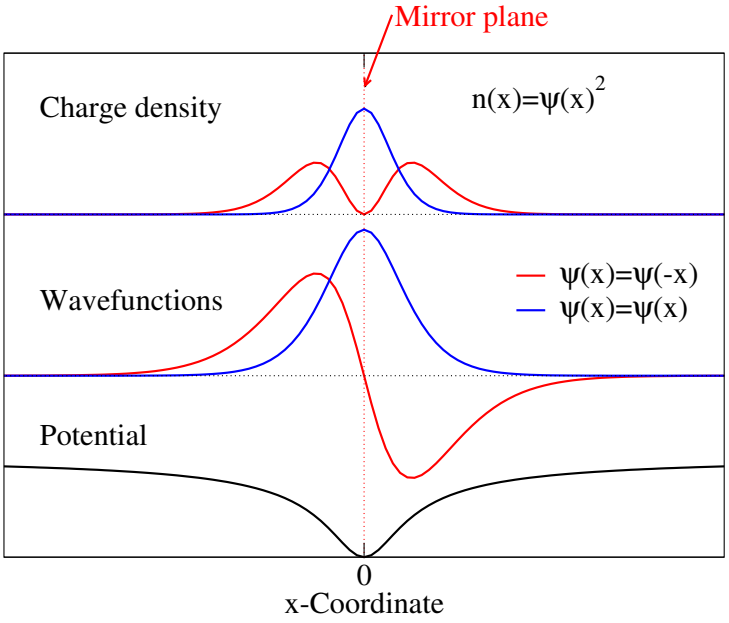


Fig. 3: *Simple example of a one-dimensional system with mirror symmetry. While the potential and the charge densities are symmetric under application of the mirror operator $x \leftrightarrow -x$ the wavefunctions can be either even or odd functions, i.e. can be symmetric or antisymmetric.*

Applying the symmetry operator twice obviously leaves the system unchanged, hence, we have $\hat{O}_M \hat{O}_M = \mathbb{1}$ and the eigenvalues $m_i = \pm 1$. These two eigenvalues correspond to wavefunctions that do not change when applying the mirror symmetry operation, i.e. wavefunctions that are 'even' for $m_i = 1$, and 'odd' functions that change their sign and have an eigenvalue of $\hat{O}_M$ of $m_i = -1$. This classification according to the eigenvalues m_i can now be used to construct an ansatz of either odd or even functions as solutions of the Hamiltonian and thus simplify the task to calculate the eigenfunctions of $\hat{H}$ substantially. Formally, this means that the functions fulfilling Eq. (24) for a particular eigenvalue $m_i = 1$ or $m_i = -1$ span disjunct eigenspaces of $\hat{H}$.

Boundary condition

When comparing solids with atoms, molecules or clusters, one main difference lies in the size of the system and the corresponding degree of localization of the wavefunction. While any real physical system is of course of finite size, this is not the only boundary condition applied in quantum mechanics. In a finite system one usually assumes that the wavefunction goes to zero on some boundary enclosing the volume in which the Hamiltonian is applied. This situation is well known from basic quantum mechanics, for example in the famous 'particle in a box' situation, and corresponds to a Hamiltonian with a discrete energy spectrum in which one can label each solution with an integer quantum number. With increasing system size the typical energy difference between the energy eigenvalues decreases. In the limit of an infinite system the eigenspectrum of the Hamiltonian turns into a function of a continuous quantum number. The probably simplest example would be the single particle in free space for which the Schrödinger equation

$$\hat{H}\psi = -\frac{1}{2}\nabla^2\psi = \epsilon\psi \tag{25}$$

is solved by a plane-wave $\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}$ and the eigenvalue is given by $\epsilon(\mathbf{k}) = \frac{1}{2}|\mathbf{k}|^2$. In the description of matter one frequently has to deal either with rather small molecules or clusters in which the wavefunction is confined to a finite space and can be assumed to be zero at some boundary enclosing the volume of interest or with systems which are so large that they can essentially be seen as infinitely large. This limit is appropriate as soon as the eigenvalues of the Hamiltonian form a continuos spectrum within the accuracy of the model.
Obviously, a general system of infinite size cannot be modeled without further assumptions. However, in solids one frequently finds an ordered arrangement of atoms which form a crystal. These crystalline solids are characterized by some building block that is repeated in all directions to generate the infinite solid. More formally, one can partition the space into a regular lattice of unit cells. The unit cell is defined as the volume spanned by three fundamental lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$. Alternatively, the three lattice parameters $|\mathbf{a}_1|$ and the angles between the vectors can be used to specify the unit-cell (the additional three parameters used if the full vectors are given just fix the rotation of the crystal in space). Each unit cell might contain several atoms of the same or of different elements at positions $\mathbf{x}_i$. The complete solid is then formed out of atoms places at positions

$$n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 + \mathbf{x}_i,$$

where $n_{1,2,3}$ denote integer numbers. Depending on the shape of the unit-cell, several different lattice systems can be defined (See e.g. [5]), together with the centering of the atoms within each unit-cell one can define the 14 different Bravais lattices. Taking symmetry operations into account that leave the origin unchanged, such as rotations or reflections, one can classify the lattice systems further into 32 crystal classes[5].

Periodic systems

The periodicity of an infinite crystal has important consequences that can be exploited in its quantum mechanical description: when the whole solid is shifted by a lattice vector

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \tag{26}$$

the resulting solid is unmodified. In our single particle description this means that $V(\mathbf{r}+\mathbf{R}) = V(\mathbf{r})$. This can be expressed by a translation operator $\hat{T}_{\mathbf{R}}$, which does not modify the Hamiltonian of the system and thus commutes with the Hamiltonian. Hence, both operators have a common set of eigenvectors and in electronic band theory one consequently determines eigenvectors ψ corresponding to the two simultaneous equations

$$\hat{H}\psi = \epsilon\psi \tag{27}$$

and

$$\hat{T}_{\mathbf{R}}\psi = \lambda\psi. \tag{28}$$

At this point an important difference arises to the discussion of the mirror symmetry: while only two eigenvalues $m_i = \pm 1$ are possible in the case of a mirror symmetry, the eigenvalue λ is in general a complex number. A restriction arises from the requirement that the wavefunction does not diverge in some direction of the infinite crystal and thus the eigenvalue must satisfy $|\lambda| = 1$, i.e. the wavefunction ψ acquires a phase factor upon translation through the crystal. For convenience one usually does not discuss the eigenvalues λ directly and each translation

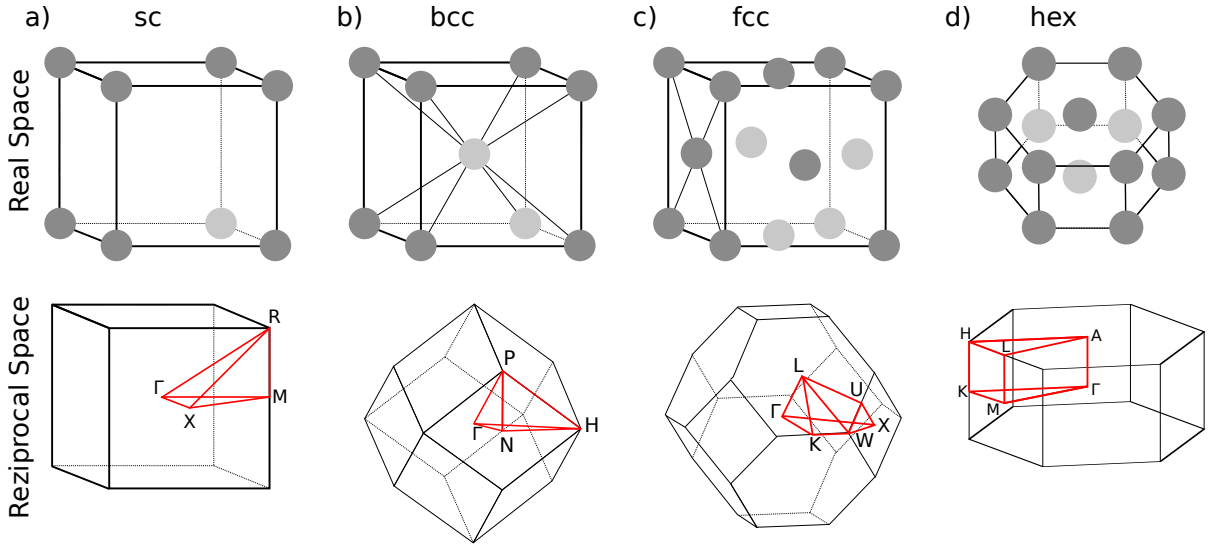


Fig. 4: *Some simple crystal structure: a) simple cubic (sc) lattice, b) body centered cubic (bcc) lattice, c) face centered cubic (fcc) lattice, d) hexagonal (hex) lattice. While the first row shows the atomic arrangement and unit cell in real space, the second column depicts the reciprocal space and the Brillouin zone. Special high-symmetry points and lines are indicated in the BZ. These lines also define the IBZ.*

operator separately but writes $\lambda = e^{i\mathbf{kR}}$ as a so-called Bloch factor. Using the corresponding Bloch ansatz

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}} u_{\mathbf{k}}(\mathbf{r}),$$

where the function $u_{\mathbf{k}}(\mathbf{r})$ is chosen to be periodic within a single unit cell (i.e. fulfilling Eq. (28) with $\lambda = 1$ for all $\mathbf{R}$), one obtains Bloch states that are eigenstates of $\hat{T}_{\mathbf{R}}$ for all lattice vectors $\mathbf{R}$ simultaneously. The power of the ansatz of Bloch states to solve Eqs. (27) and (28) can be understood when considering the Schröderinger equation and applying the product rule for the differential operator

$$\left(-\frac{1}{2}\nabla^2 + V(\mathbf{r})\right) e^{i\mathbf{kr}} u_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}} \left(-\frac{1}{2}(\nabla + i\mathbf{k})^2 + V(\mathbf{r})\right) u_{\mathbf{k}}(\mathbf{r}) = \epsilon_{\mathbf{k}} e^{i\mathbf{kr}} u_{\mathbf{k}}(\mathbf{r}) \qquad (29)$$

which turns into an equation for $u_{\mathbf{k}}(\mathbf{r})$ after multiplication with $e^{-i\mathbf{kr}}$. This equation for $u_{\mathbf{k}}$ has to be solved only inside a single unit cell with periodic boundary conditions being applied. The use of eigenfunctions of the translation operator to solve for eigenfunctions of the Hamiltonian thus reduces the seemingly impossible task of treating the Hamiltonian of the infinite crystal to an eigenvalue problem inside a single unit cell. However, one now has to solve separate eigenproblem for different $\mathbf{k}$-values. Nevertheless, Eq. (29) represents a tremendous simplification and makes it possible to treat periodic structures very efficiently.

Brillouin zone

We introduced the $\mathbf{k}$-vector as an efficient way to construct wavefunctions that are eigenfunctions of the translation operators $\hat{T}_{\mathbf{R}}$. Since the lattice vectors are given as linear combinations of the fundamental lattice vectors with integer coefficients, see Eq. (26), only a restricted set of

$\mathbf{k}$ values has to be considered. This is most easily seen in a one-dimensional example. For a lattice constant a, the $k-$values

$$k = k_0 + m\frac{2\pi}{a}$$

with arbitrary integer m correspond to the same eigenvalue λ for all lattice vectors $R = na$ because

$$\lambda = e^{ikR} = e^{i(k_0+m\frac{2\pi}{a})(na)} = e^{ik_0a}.$$

Hence, it is sufficient to restrict the k value to an interval $-\frac{\pi}{a} \ldots \frac{\pi}{a}$ due to the periodicity of the Bloch factors. Similarly, in three dimensions the $\mathbf{k}$-vector can be restricted to a volume spanned by the three vectors

$$\mathbf{b}_1 = \frac{2\pi}{V}\left(\mathbf{a}_2 \times \mathbf{a}_3\right), \quad \mathbf{b}_2 = \frac{2\pi}{V}\left(\mathbf{a}_3 \times \mathbf{a}_1\right), \quad \mathbf{b}_3 = \frac{2\pi}{V}\left(\mathbf{a}_1 \times \mathbf{a}_2\right),$$

where V denotes the volume of the unit cell of the crystal. In complete analogy to the crystal lattice which is spanned by the lattice vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, these vectors span the so-called reciprocal lattice. The unit cell of the reciprocal lattice, from which the $\mathbf{k}$-vectors can be chosen is called the (first) Brillouin zone (BZ). Frequently, additional symmetries of the system can be used to relate solutions at different $\mathbf{k}$-values inside the BZ such that only a part of the BZ actually has to be considered to obtain all solutions of Eqs. (27) and (28). This part is then called the irreducible Brillouin zone (IBZ). The Brillouin zones for a few simple Bravais lattices are visualized in Fig. 4.

3.2 The bandstructure

The eigenvalues $\epsilon(\mathbf{k})$ plotted against the k-points along some lines constitute a bandstructure plot, a basic tool to visualize the electronic structure of the system. Examples of bandstructures of two simple materials are shown in Fig. 5. At each of the $\mathbf{k}$-values an infinite set of solutions of Eq. (27) can be constructed. For the different $\mathbf{k}$-points these form continuos bands as can be seen by the continuous change of the effective $\mathbf{k}$-dependent Hamiltonian of Eq. (29). One should realize, that these "bands" are in fact cuts through surfaces in the three-dimensional reciprocal space and thus always only show a small part of the electronic structure.
On the other hand, one usually shows the bandstructure along lines of high symmetry, i.e. for $\mathbf{k}$-vectors which are mapped by a symmetry operation on itself or on a $\mathbf{k}$-vector that is shifted by a reciprocal lattice vector. The reasoning behind this choice is that these point and lines of high symmetry are frequently extrema of the bandstructure. For example, if one considers the so called Γ-point $\mathbf{k} = (0,0,0)$ and a system with inversion symmetry, one has $\epsilon(\mathbf{k}) = \epsilon(-\mathbf{k})$. If now there is only a single band involved in this discussion $\epsilon(\mathbf{k})$ must have an extrema at $\mathbf{k} = (0,0,0)$. For the different lattice structures and symmetries different high-symmetry points exist. Fig. 4 shows simple Brillouin zones with their symmetry points and conventional naming. Actually, the a system does not need to have spatial inversion symmetry in order to have a bandstructure with $\epsilon(\mathbf{k}) = \epsilon(-\mathbf{k})$. From time reversal symmetry one deduces that with $\psi_\mathbf{k}$ also $\psi^*_\mathbf{k}$ is an eigenstate of the Hamiltonian with the same eigenvalue. Furthermore, from Eq. (29) one can deduce that $u_\mathbf{k} = u^*_{-\mathbf{k}}$ and thus $\psi^*_\mathbf{k} = \psi_{-\mathbf{k}}$ and therefore one finds $\epsilon(\mathbf{k}) = \epsilon(-\mathbf{k})$. A more careful analysis of time reversal symmetry is actually needed if one also considers the spin degree of freedom. As time-inversion changes the direction of the spin one finds in this case $\epsilon_\sigma(\mathbf{k}) = \epsilon_{-\sigma}(-\mathbf{k})$. In the case of a magnetic system, i.e. a system in which one finds different effective potentials for electrons of different spins, one additionally has to take

care that the magnetization direction encoded into the potential also changes sign under time-reversal symmetry. For a fixed magnetization direction or in other words for a fixed $\mathbf{B}$-field the Hamiltonian no longer shows time reversal symmetry.

Complex bandstructure

While the procedure to choose a $\mathbf{k}-$point at which the Hamiltonian is then diagonalized is the usual way to view Eqs. (27) and (28), in principle both equations can be viewed on the same level and one can also solve for $\mathbf{k}(\epsilon)$. This is often called an inverse problem but mathematically there is little difference between both equations. However, the physical significance of the quantities impose some restrictions. First, the eigenvalue of the Hamiltonian, interpreted as the energy, must be real. While one could argue that this should be automatically be the case as the Hamiltonian is hermitian, this hermicity is actually tied to the boundary condition we impose on the wavefunction and in general only periodic boundary conditions as ensured by Eq. (28) with a real $\mathbf{k}$ will lead to a hermitian eigenvalue problem in Eq. (27). This restriction to real $\mathbf{k}$-points is of course necessary if one wants to calculate wavefunctions of an infinite solid as for any imaginary part $\boldsymbol{\kappa}$ of $\mathbf{k} = \mathbf{q} + i\boldsymbol{\kappa}$ one obtains a wavefunction that grows/decays exponentially in the direction of $\boldsymbol{\kappa}$. Such a wavefunction can not be normalized in any infinite system. However, there are situations in which the crystal is actually finite, i.e. a system where the periodic potential stops at some point. Most notable examples are surfaces of so called semi-infinite systems or -more general- interfaces in which two of such semi-infinite half-spaces are matched. In this case the exponential growth of the wavefunction will be stopped at this interface and the wavefunction decaying away from the interface is a completely physical solution of Eqs. (27) and (28). To obtain these solutions one has to solve Eqs. (27) and (28) at fixed real energy ϵ for the full spectrum of $\hat{T}_{\mathbf{R}}$ with the eigenvalues λ containing all usual states with real values of $\mathbf{k}$ ($|\lambda| = 1$) as well as so called evanescent states with $|\lambda| < 1$. Similar to the infinitely many solutions of Eq. (27) for fixed $\mathbf{k}$, i.e. for the infinitely many bands in an infinite crystal, one can find infinitely many eigenvalues of $T_{\mathbf{R}}$ at fixed energy, nearly all of which correspond to evanescent states[6, 7].
The most basic example of such an evanescent state is known from the elementary quantum physics problem of tunneling through a rectangular potential barrier. This problem can be solved by matching the different solutions of the Hamiltonian by value and slope, i.e. simple plane-waves outside the barrier and functions behaving like $e^{\pm\kappa r}$ inside the barrier. This already demonstrates the situations in which these evanescent states have physical significance. They describe the decay of the wavefunctions in situations in which no usual Bloch state exist at the energy in question. In particular, the solution of Eqs. (27) and (28) for $\mathbf{k} = \mathbf{q} + i\boldsymbol{\kappa}$ also describes the decaying states in the potential contained in the Hamiltonian. Fig. 6 shows a plot of the bandstructure of $SrTiO_3$ containing not only the Bloch states but also these evanescent bands.

The Fermi energy

While the fermionic character of the electrons leads to the well known anti-symmetry relations of the many-body wavefunction, the description in terms of single-electrons discussed here takes the Pauli principle into account by requiring that each single electron state is occupied by at most one electron. At the same time the ground state of the system, i.e. the state with minimal energy is naturally assumed to be that configuration in which the lowest energy single-particle

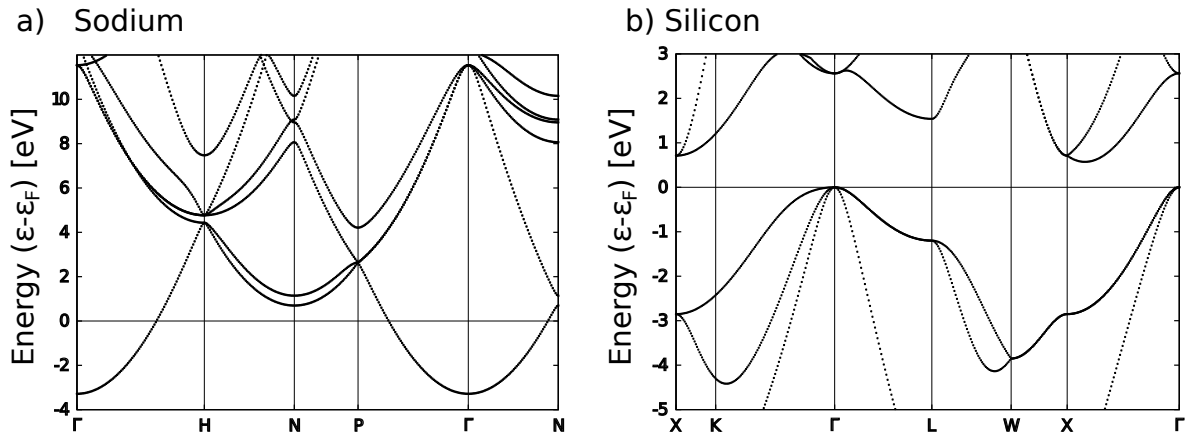


Fig. 5: *Bandstructure of Silicon. The different bands are plotted for k-points along some high symmetry lines.*

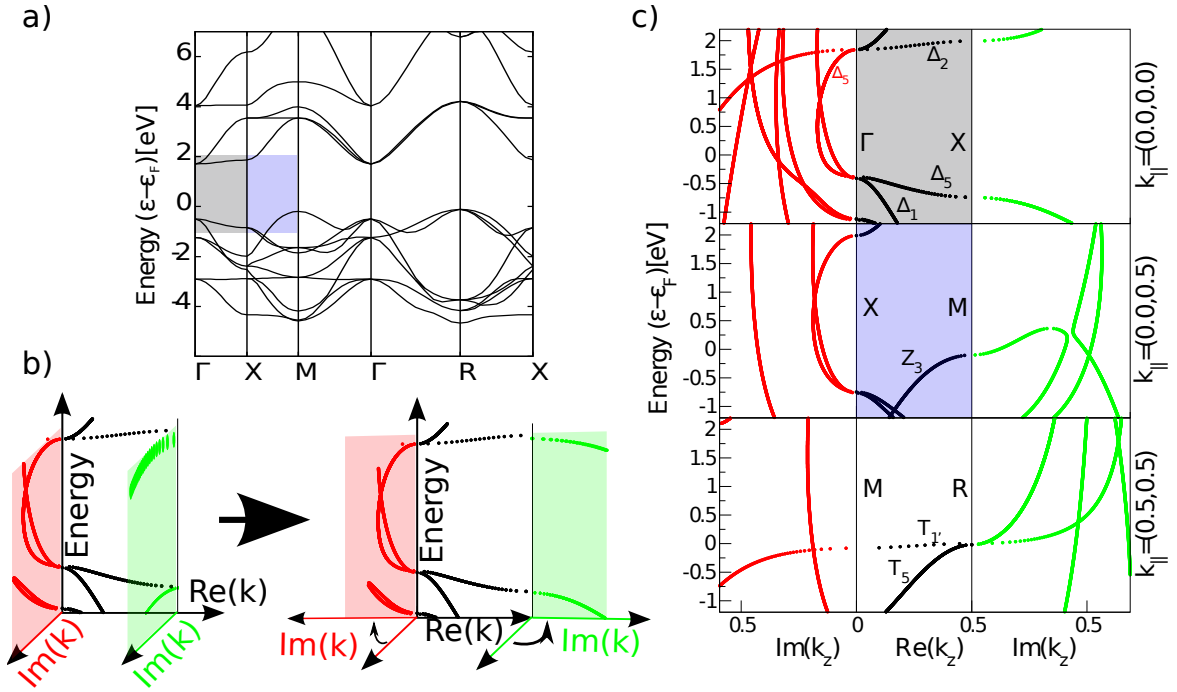


Fig. 6: *Complex bandstructure of $SrTiO_3$. a) Usual bandstructure (BS) of the transition metal oxide in the cubic perovskite structure. b) Schematics of the construction of the plot of the complex BS. Additional bands with complex* **k***-values appear at right angle in complex* **k***-space at the high-symmetry points. These bands in the red and green panel are then plotted adjacent to the Bloch states by rotating these $Im(k), \epsilon$ panels next to the $Re(k), \epsilon$ plot. c) complex BS for some special lines in* **k***-space. The parts of the real Bloch BS which is also shown in a) is indicated by shaded areas.*

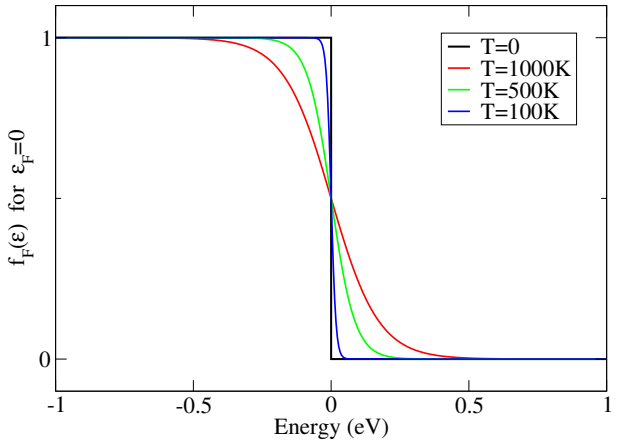


Fig. 7: *Fermi function at different temperatures. While the Fermi functions equals a sharp step-function at $T = 0$ it become increasingly smooth at higher temperatures*

states are occupied and the higher states remain empty. In detail, if there are N electrons in the system one occupies the lowest N eigenstates of the Hamiltonian such that

$$N = \frac{8\pi^3}{V} \int_{BZ} \sum_i^{\infty} f_F(\epsilon_i(\mathbf{k}))\, d^3k \approx \frac{8\pi^3}{V} \int_{BZ} \sum_i^{\epsilon_i(\mathbf{k})<\epsilon_F} 1\, d^3k. \tag{30}$$

Here one often introduces the Fermi distribution function (see Fig. 7)

$$f_F(\epsilon) = \frac{1}{1 + e^{\frac{\epsilon-\epsilon_F}{k_B T}}}$$

which includes the effect of the temperature smearing at finite temperatures T. It equals the step function at zero temperature and all states below a critical energy ϵ_F, the so called Fermi-energy, are occupied. The k-integration is performed over all of the first Brillouin zone(BZ) with proper normalization $\frac{8\pi^3}{V}$ and one has to sum over all bands ϵ_i. As a side-note, one should mention that in this presentation the index i will also contain the spin-quantum number, hence in a non-magnetic system each state will be doubly degenerate. Equation (30) then determines ϵ_F so that for example the charge density of the system can be obtained as

$$n(\mathbf{r}) = \frac{8\pi^3}{V} \int_{BZ} \sum_i^{\infty} f_F(\epsilon_i(\mathbf{k}))|\psi_{i,\mathbf{k}}(\mathbf{r})|^2\, d^3k \approx \frac{8\pi^3}{V} \int_{BZ} \sum_i^{\epsilon_i(\mathbf{k})<\epsilon_F} |\psi_{i,\mathbf{k}}(\mathbf{r})|^2 d^3k.$$

The electronic structure of the system in the vicinity of the Fermi energy is of fundamental interest. As one learns from quantum mechanical perturbation theory, the response of the system to external perturbations which might be applied in an experiment to probe its property will accompanied with an intermixing of different states ψ_i. Charge conservation and the Pauli principle then favors a response in which a former unoccupied state above the Fermi-energy interacts with an occupied state below and hence the highest occupied and lowest unoccupied states in a system determine most of the response of the system. Furthermore, the description in terms of single electron states is most appropriate close to the Fermi-energy as the single particle states near to the Fermi-level can be shown to have the largest lifetime (cf. Sec. 4.3).

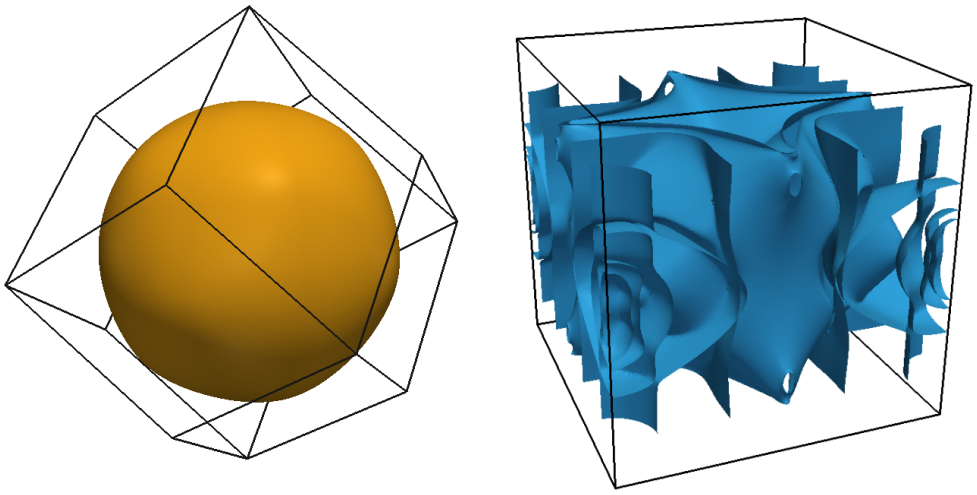


Fig. 8: *Comparison between a simple Fermi surface for the alkali metal bcc-Rb (left) to a complicated, multi-sheeted Fermi surface of the ordered FePt- alloy in the L10 -structure.*

Fermi surface

As the states directly at the Fermi level determine the low energy response of the system, it can be useful to plot the surface in $\mathbf{k}$-space defined by the equation $\epsilon(\mathbf{k}) = \epsilon_F$. This surface can consist of different disconnected sheets corresponding to the different bands cutting the Fermi level and in some cases has rather complex forms. An example of such a Fermi-surface plot with additional data superimposed can be seen on the cover of this book and another example is shown in Fig. 8. Of special interest are regions in which the Fermi surface shows areas in which the different parts of the surfaces are parallel to each other as this indicates the existence of the special $\mathbf{k}$-vector, i.e. a special periodicity, for which the system is particularly susceptible for perturbations. The enhanced response or instability of the system is then assigned to this so called Fermi-surface nesting effect.

Metals and Insulators

One can now distinguish two cases: either there is an electronic band intersecting the Fermi-energy, i.e. there are states at, directly above and directly below the Fermi-level or the Fermi-energy lies in a region where there is a gap in the bandstructure. In the first case one deals with a metallic material while in the case with an electronic gap at the Fermi-level one talks of an insulator or, if the gap is small, of a semiconductor. The obvious consequence of the presence of a gap in the bandstructure is the need to invest a finite amount of energy to promote an electron from an occupied state into an unoccupied state. Hence, this transition can not be achieved by small perturbing electric fields, i.e. electrons can not change their electronic state and no electrical current flows in response to a small applied bias voltage to the material, the very definition of an insulating material.

Velocity operator

Of course all physical quantities can be expressed in terms of the eigenfunctions and expectation values of the corresponding operators. However some properties of the system can also be

obtained from the bandstructure alone. For example, the mean velocity of the electrons in a Bloch state can be evaluate from the bandstructure. To show this one starts by expanding

$$\epsilon(\mathbf{k}) \approx \epsilon(\mathbf{k}_0) + \nabla_{\mathbf{k}}\epsilon(\mathbf{k})|_{k=k_0}(\mathbf{k}-\mathbf{k}_0)$$

up to first order around $\mathbf{k}_0$. The same expansion can also be generated by using the corresponding expansion of the $\mathbf{k}$-dependent Hamiltonian for the Bloch functions $u_{\mathbf{k}}$ (Eq. (29))

$$H_{\mathbf{k}} \approx H_{\mathbf{k}_0} + (\mathbf{k}-\mathbf{k}_0)\,(\mathbf{k}_0 - i\nabla)$$

together with simple first order perturbation theory and the product rule for the Bloch states. By comparing these expansions one obtains for the first order expansion coefficient

$$\nabla_{\mathbf{k}}\epsilon = -i\int d\mathbf{r}\psi^*_{\mathbf{k}}(\mathbf{r})\nabla_{\mathbf{r}}\psi_{\mathbf{k}}(\mathbf{r}).$$

As $\mathbf{p}/m = -i\nabla_{\mathbf{r}}/m$ is the velocity operator, one can identify the derivative of the bandstructure with the velocity of the electron

$$\mathbf{v} = \nabla_{\mathbf{k}}\epsilon(\mathbf{k}) = \frac{\partial\epsilon(\mathbf{k})}{\partial\mathbf{k}}.$$

For small velocities, i.e. around the minima/maxima of the bands where $\nabla_{\mathbf{k}}\epsilon = 0$ the bandstructure is parabolic in shape and one can draw an analogy to free particles. As in this case

$$\epsilon_{free} = \frac{1}{2m}p^2$$

with the mass m of the electron. To define an analogous equation in the case of an infinite crystal one might recall that the momentum $\mathbf{p}$ is actually related to the translational invariance of the free particle. More formally, the phase λ acquired by the wavefunction upon translation by $\mathbf{r}$ is given by

$$\hat{T}_{\mathbf{r}}\psi = e^{i\mathbf{pr}}\psi$$

as $\lambda = e^{i\mathbf{pr}}$. The formal equivalence of this equation with Eq. (28) suggests to interpret $\mathbf{k}$ as a kind of momentum vector. To clarify the difference to the usual momentum vector, one speaks of a "crystal momentum". To complete the analogy, one can use the expansion up to second order around minima/maxima of the bandstructure (where for simplicity the value of the extremum is set to zero as well the $\mathbf{k}$-value at which it occurs)

$$\epsilon(\mathbf{k}) = \sum_{ij}\frac{\partial^2\epsilon}{\partial k_i\partial k_j}k_ik_j$$

to define an "effective mass" of the Bloch electron as

$$\tilde{m}_{ij} = \frac{1}{2}\left(\frac{\partial^2\epsilon}{\partial k_i\partial k_j}\right)^{-1}.$$

The effective mass in this definition is actually a tensor quantity as it depends on the directions in which the derivatives are taken.

Density of states

While the discussion of the bandstructure, i.e. of the eigenvalues and the corresponding eigenfunctions enables a complete description of the electronic structure, it is often more appropriate to think about energy dependent quantities. In particular, many spectroscopic experimental methods determine the energy dependent response of the system making a interpretation of such experiments in terms of functions of energy very natural. The most basic of such concepts is the density of states (DOS), which directly related to the bandstructure. It counts the number of states available at some energy. This idea can easily be cast into an equation using the Dirac delta function

$$n_{\mathrm{DOS}}(\epsilon) = \frac{8\pi^3}{V} \int_{BZ} \sum_i^{\infty} \delta(\epsilon_i(\mathbf{k}) - \epsilon)\, d^3k. \tag{31}$$

An alternative equivalent definition can be given in terms of integrals over constant energy surfaces in k-space

$$n_{\mathrm{DOS}}(\epsilon) = \sum_i^{\infty} \int_{\epsilon_i(\mathbf{k})=\epsilon} \frac{d^2k}{4\pi^3} \frac{1}{|\nabla_{\mathbf{k}} \epsilon_i(\mathbf{k})|}. \tag{32}$$

While Eq. (31) can be defined for any setup, the Eq. (32) is only valid for infinite periodic systems. Both of these definitions already show some basic features of the DOS. Whenever the bandstructure if "flat", i.e. $\epsilon(\mathbf{k})$ changes only little with energy, one has a big number of states within any given energy range and the DOS is large. In contrast, at energies at which there is no energy eigenstate $\epsilon(\mathbf{k})$, i.e. within bandgaps, the DOS is zero. Furthermore, denoting the number of states present at some energy, the DOS is of course always positive and the integral

$$N = \int_{\infty}^{\epsilon_F} n_{\mathrm{DOS}}(\epsilon) d\epsilon \tag{33}$$

again gives the total number of electrons in the system. From Eq. (31) one also finds a fundamental difference between finite systems, in which the DOS is just a sum of delta functions and an infinite system in which the spacing between these delta functions goes to zero and the DOS turns into a continuos function of energy. Most challenging to study are often situations in which both kind of features are present in the same systems as it can happen if localized "impurity" states are present in an infinite environment.
In plots showing the density of states one frequently uses the weight of the states in some region of space to produce different projections of the DOS. In particular this can be used to assign specific states to different atoms or investigate the symmetry of states. To understand this technique one can use the eigenfunction, or better the probability distribution due to the square of the eigenfunctions, to generalize the density of states to the spectral function

$$A(\mathbf{r}, \epsilon) = \frac{8\pi^3}{V} \int_{BZ} \sum_i^{\infty} \delta(\epsilon_i(\mathbf{k}) - \epsilon)\, |\psi_{i,\mathbf{k}}(\mathbf{r})|^2\, d^3k. \tag{34}$$

Using this definition the total charge density is given as an energy integral over the spectral function as a generalization of Eq. (33) to

$$n(\mathbf{r}) = \int_{-\infty}^{\epsilon_F} A(\mathbf{r}, \epsilon) d\epsilon \tag{35}$$

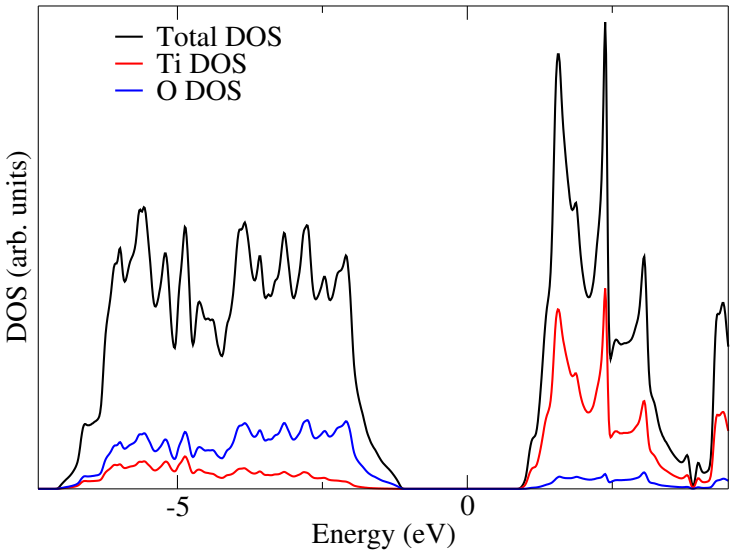


Fig. 9: *Density of states of TiO_2 in the rutile structure. The different lines show the total DOS and the projections on spheres around the different atomic nuclei. The remaining contribution mostly come from outside the projection spheres.*

and the density of states is the spatial average of the spectral function

$$n_{\mathrm{DOS}}(\epsilon) = \int_V A(\mathbf{r}, \epsilon) dr^3. \tag{36}$$

This definition now allows to define a local density of states LDOS, in which the integration volume V might be restricted to some area of interest such as the vicinity of a specific atom or the surface layer of a crystal. An example of such a decomposed density of states is shown in Fig. 9. Other decompositions of the local density of states usually called orbital decomposed (L)DOS additionally expands the wavefunction into spherical harmonics around some atom and then uses only specific $l-$values to generate the spectral function and the LDOS (see Fig. 10). This orbital decomposed DOS allows the description of the system in terms of an atomic picture, in which one speaks about $s, p, d, f-$states and actually refers to Bloch states which have mainly such an s, p, d, f orbital character close to some atom. While such a discussion can be very useful and appropriate one should keep in mind its origin and limitations.

3.3 Wannier functions

So far we discussed the description of the infinite solid in terms of Bloch states, which are by construction extending over all space. However, in many situations it is more appropriate to use a description in terms of localized functions. In particular the description in terms of functions localized at the different atomic sites often provides a complementary picture of the system. While many of such localized functions exist which can be used as a basis set to expand the wavefunctions there is also a specific class of functions derived from the Bloch states: the Wannier functions

$$w_{\mathbf{R}}(\mathbf{r}) = \frac{1}{V_{BZ}} \int_{BZ} d^3k \, \psi_{\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{R}}. \tag{37}$$

These functions have a set of useful properties:

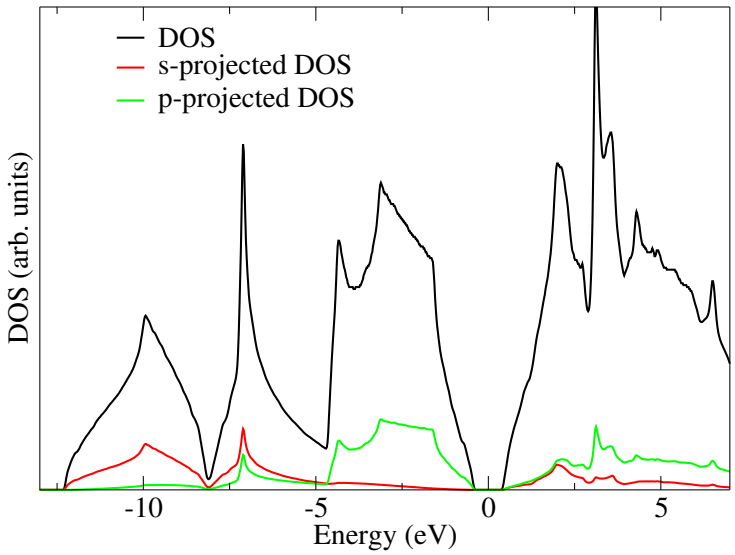


Fig. 10: *Density of states of Silicon. The different lines show the total DOS and the projections on the $l = 0, 1$ or $l = s, p$ wavefunctions on spheres around the atomic nuclei. The remaining contribution mostly come from outside the projection spheres.*

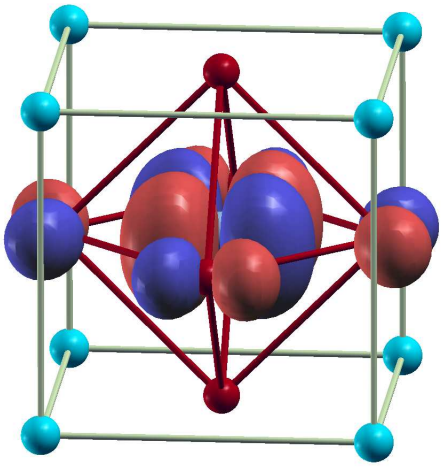


Fig. 11: *Wannier function in a perovskite oxide material. The Wannier function is localized at the central metal atom and shows the typical shape of a d_{xy}-orbital at this atom and the shape of a p-orbital at nearby oxygen atoms.*

- The Wannier functions for different lattice vectors $\mathbf{R}$ are the same except for a simple displacement, i.e. the Wannier function actually depends only on $\mathbf{r}-\mathbf{R}$ and one can write $w_{\mathbf{R}}(\mathbf{r}) = w(\mathbf{r}-\mathbf{R})$.

- They are orthonormal functions due to the orthonormality of the underlying Bloch functions

$$\begin{aligned} \langle w_{\mathbf{R}}|w_{\mathbf{R}'}\rangle &= \int d^3r w^*_{\mathbf{R}}(\mathbf{r}) w_{\mathbf{R}'}(\mathbf{r}) && (38)\\ &= \frac{1}{V_{BZ}^2}\int_{BZ} d^3k \int_{BZ} d^3k' \int d^3r \psi^*_{\mathbf{k}}(\mathbf{r})\psi_{\mathbf{k}'}(\mathbf{r}) e^{i\mathbf{k}\mathbf{R}} e^{-i\mathbf{k}'\mathbf{R}'} && (39)\\ &= \delta(\mathbf{R}-\mathbf{R}'). && (40) \end{aligned}$$

- Wannier functions are not defined uniquely. This is a consequence of the fact that the phase of the Bloch states $\psi_{\mathbf{k}}(\mathbf{r})$ is not defined and can be chosen freely. Hence Eq. (37) can be modified by multiplying each Bloch state by an arbitrary phase factor $c_{\mathbf{k}}$ to

$$w_{\mathbf{R}}(\mathbf{r}) = \frac{1}{V_{BZ}}\int_{BZ} d^3k\, c_{\mathbf{k}} \psi_{\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{R}}. \quad (41)$$

  This freedom can be used to construct Wannier functions with special properties, most notable to construct localized Wannier functions. In the scheme of Marzari and Vanderbilt [8] the phase-factors are chosen such that the spread of the Wannier functions is minimized. While these so called maximally localized Wannier functions are in many cases the most efficient set of Wannier functions to use their calculation can also be quite cumbersome and therefore other simpler schemes to construct localized Wannier functions are also in use. One can show that these Wannier functions can be localized such that they decay exponentially [9]. The situation is more complicated in cases in which the individual bands are not identified easily, i.e. if there are many solutions $\epsilon(i,\mathbf{k})$ where the different bands are labeled by i. In these situations the transformation between Wannier functions and Bloch functions is generalized to

$$w_{i,\mathbf{R}}(\mathbf{r}) = \frac{1}{V_{BZ}}\sum_j \int_{BZ} d^3k\, U_{i,j}(\mathbf{k}) \psi_{j,\mathbf{k}}(\mathbf{r}) e^{-i\mathbf{k}\mathbf{R}}, \quad (42)$$

  i.e. the simple phase factor $c_{\mathbf{k}}$ is replaces by a unitary transformation matrix $U_{i,j}(\mathbf{k})$. Sophisticated algorithms have been developed in this situation to deal with the entangled bandstructure and to obtain localized Wannier functions in this case[10, 11].

- From the Wannier functions Bloch states can be reconstructed as

$$\psi_{i,\mathbf{k}}(\mathbf{r}) = \frac{1}{N}\sum_j \sum_{\mathbf{R}} e^{i\mathbf{k}\mathbf{R}} U^*_{i,j}(\mathbf{k}) w_{j,\mathbf{R}}(\mathbf{r}). \quad (43)$$

  This relation (together with Eq. (42)) shows that the Bloch states and the Wannier functions span the same function space. In other words, these two sets of functions are related by a simple unitary transformation.

This last property is the key to many applications of Wannier functions. Since they describe the same states as the Bloch states of the band, localized Wannier functions provide the natural tool to construct models in which a description of the system in terms of localized functions is required, i.e. Wannier functions can be used to transform the Hamiltonian of the system into a tight-binding representation. By constructing Wannier functions out of the single (or few bands) of interest for the model one can effectively split the system into a part described in terms of Bloch bands and a subset of states treated by the model. Such an approach is often employed to link models treating the complex interaction of the electrons in the band, e.g. in the Hubbard model, with a material specific description of the bandstructure. More formally one can consider the Wannier functions as a basis set in which the Hamiltonian is no longer diagonal, e.g. the Wannier functions are not eigenfunctions, but block-diagonal where the different blocks correspond to different bands. One then replaces one of these blocks by a different model. This basic idea of applying model Hamiltonians only to a subset of relevant states is the principle used in many schemes we will discuss in this school and Wannier functions can be used to provide a defined interface between these models and an effective single particle description.
Another useful application of Wannier functions is the so called Wannier interpolation scheme. It allows to calculate a quantity $a(\mathbf{k})$, i.e. any $\mathbf{k}$ dependent quantity (or integrals of these), on a relatively coarse grid of $\mathbf{k}$-points and then to accurately interpolate the results from these $\mathbf{k}$-points to a much denser grid. Its basic idea is simple and can easily be understood in terms of the bandstructure. Using the Dirac notation for the matrix elements and taking into account that the Bloch functions diagonalize the Hamiltonian, the eigenvalues $\epsilon(i,\mathbf{k})$ are given as

$$\begin{aligned}
\epsilon(i,\mathbf{k}) &= \langle\psi_{i,\mathbf{k}}|H|\psi_{i,\mathbf{k}}\rangle \\
&= \sum_{n,n'}\sum_{\mathbf{R}}\sum_{\mathbf{R}'} U_{i,n}(\mathbf{k})\langle w_{n,\mathbf{R}}|\hat{H}|w_{n',\mathbf{R}'}\rangle U^*_{n',i}(\mathbf{k})e^{i\mathbf{k}(\mathbf{R}-\mathbf{R}')} \\
&= \sum_{n,n'}\sum_{\mathbf{R}} U_{i,n}(\mathbf{k})\langle w_{n,\mathbf{R}}|\hat{H}|w_{n',\mathbf{0}}\rangle e^{i\mathbf{k}\mathbf{R}} U^*_{n',i}(\mathbf{k}) \\
&= diag\left(\sum_{\mathbf{R}}\langle w_{n,\mathbf{R}}|\hat{H}|w_{n',\mathbf{0}}\rangle e^{i\mathbf{k}\mathbf{R}}\right)_i ,
\end{aligned}$$

where first Eq. (42) was used, then the translational property of the Wannier functions has been exploited and in the last step the application of the matrices U are understood as a simple diagonalization. Hence, one calculates the Wannier functions from some coarse grid of $\mathbf{k}$ values and evaluates the Hamiltonian matrix elements in this basis, then one can evaluate for any $\mathbf{k}$-value the bandstructure $\epsilon(i,\mathbf{k})$ by a simple diagonalization of a small matrix. The size of the matrix is given by the number of bands one considers in the construction of the Wannier functions. The power of this approximation lies in the fact that the matrix elements

$$\langle w_{n,\mathbf{R}}|H|w_{n',\mathbf{0}}\rangle$$

decay rapidly with increasing $\mathbf{R}$ due to the localization of the Wannier functions and thus only a few $\mathbf{R}$ and corresponding $\mathbf{k}$ values are needed to obtain a good approximation of the matrix.
Finally, one should be aware of the fact that Wannier functions also provide a powerful tool to calculate quantities that are difficult to access in the Bloch description. For example, as the position operator $\hat{\mathbf{r}}$ is ill-defined in infinite periodic systems, quantities like the electric polarization given by the operator $\hat{\mathbf{P}} = e\hat{\mathbf{r}}$ cannot be defined easily. It can be shown however, that the changes of such quantities can be expressed in terms of Wannier functions. This leads

to the so called modern theory of polarization which is equivalent to a description in terms of the Berry phase [10]. Similarly, quantities like orbital moments or responses to electrical fields can be treated.

4 Approximations to the many-electron problem

So far we considered only the single-particle form of the Hamiltonian and the question arises what relevance this discussion can have. Before eluding on that question we first would like to illustrate that aiming at an exact and complete solution of the many-electron Schrödinger equation is a pointless enterprise.

Upon inclusion of the interaction term, $\hat{V}_{ee} \neq 0$, the Hamiltonian Eq. (5) does not separate into single-particle terms anymore, i.e. it cannot be written as a sum of single-particle Hamiltonians $\hat{h}$. Consequently, the resulting many-body Schrödinger equation (13) is extremely difficult to solve for any reasonable number of electrons. Even if we could solve this equation, the resulting wavefunction $\Psi(\mathbf{r}_1 ... \mathbf{r}_N)$ is a very complicated object. To illustrate this point we consider a single nitrogen atom which has seven electrons. Hence, we need to calculate $\Psi(\mathbf{r}_1 ... \mathbf{r}_7)$ which depends on $7 \cdot 3 = 21$ one-dimensional coordinates. In order to make a rough table of the values of Ψ at different positions of the seven electrons we use an extremely coarse grid with only ten points in each variable. Our table will then have 10^{21} entries which, assuming a memory need of only one byte per entry, corresponds to a memory requirement of 10^{21} bytes for storing the wavefunction in a computer. Using a micro SD card with $64\ \mathrm{GB} = 6.4 \times 10^{10}$ bytes of storage space, we need 1.56×10^{10} micro SD cards. Even if each of these cards only weights one gram their combined mass is 1.56×10^4 tons. This discussion shows that even for an extremely simple system with only seven electrons and a very rough numerical approximation for the wavefunction we would need around 500 trucks to transport those SD cards.

In addition, we are not really interested in the wavefunction itself. In quantum mechanics, the wavefunction is merely a tool in the calculation of observables which are given as the expectation values of the corresponding operator with respect to the wavefunction, i.e. for a general operator $\hat{O}$

$$\langle \hat{O} \rangle = \langle \Psi | \hat{O} | \Psi \rangle. \tag{44}$$

In other words, we would be making an extreme effort to calculate the wavefunction just to throw away most of the information contained in it directly afterwards. Ideally, one would calculate the expectation values of any relevant operator with respect to the many-body wavefunction without calculating the exact wavefunction itself.

This can be achieved by either calculating an approximate wavefunction with which one calculates the expectation values, by finding a different way to calculate expectation values or by approximating the many-body Hamiltonian. The first idea leads to quantum chemical methods like configuration interaction (CI) or coupled cluster expansions while density functional theory or Greens function methods fall in the second category. While these are rather general methods, approximations to the Hamiltonian are chosen such that the Hamiltonian contains the relevant physics. The key in such an approach is to find the balance between including all effects one wants to study correctly and still obtaining a model Hamiltonian and Schrödinger equation that can be solved, e.g. with renormalization group techniques.

4.1 Approximations to the wavefunction

Our simple example of the gigantic data needed to store the wavefunction of the nitrogen atom already indicates a possible strategy to tackle the many-electron problem. Instead of trying to put the wavefunction on a numerical grid in space, one needs to find a more efficient representation of the wavefunction. A systematic approach can be developed by starting at the non-interacting system and by recalling that the many-electron wavefunction in this case was a Slater determinant.
The many-body wavefunction of a system of interacting electrons cannot be written as a single Slater determinant since the Hamiltonian does not separate into single-particle terms in this case. However, one can use the Slater determinants as a basis to expand the many-body wavefunction, i.e. the many-body wavefunction is given as

$$\Psi(\mathbf{x}_1 \dots \mathbf{x}_N) = \sum_j c_j \Psi_j^{\text{Slater}}(\mathbf{x}_1 \dots \mathbf{x}_N), \tag{45}$$

where Ψ_j^{Slater} denote Slater determinants containing different sets of single-particle orbitals for different j. In order to preserve the normalization of the wavefunction the coefficients have to satisfy $\sum |c_j|^2 = 1$. It can be shown that the Slater determinants form a complete set of basis functions provided that the single-particle orbitals $\psi_i(\mathbf{x})$ which are used to construct the determinants, see Eq. (20), form a complete single-particle basis. Truncating the expansion (45) after the first term, i.e. approximating the many-body wavefunction as a single Slater determinant despite of the interaction, leads to the Hartree-Fock approximation. The Hartree-Fock orbitals are then determined by minimizing the total energy with respect to the single-particle orbitals $\psi(\mathbf{r})$. In other words, the Hartree-Fock Slater determinant is the variationally best single Slater determinant approximation to the true many-body wavefunction. Naturally, this approximation is valid if the true many-body wavefunction is dominated by one Slater determinant. If the expansion Eq. (45) contains several coefficients of approximately equal weight the Hartree-Fock approximation will break down. In the configuration interaction method one uses the Hartree-Fock orbitals as a single-particle basis and includes more determinants which are labeled by the number of orbitals that are changed compared to the Hartree-Fock determinant. For example, in CI singles one includes all Slater determinants where a single orbital from the Hartree-Fock determinant is replaced by an orbital that is unoccupied in the Hartree-Fock method. One can then optimize the coefficients c_j to find the best representation of the many-body wavefunction in this set of Slater determinants. Adding more terms to the expansion, for example replacing two or three orbitals, leads to a complicated and expensive computational calculations which yield high accuracy but can be performed only for a small number of electrons N. For a more detailed description of methods based on approximations of the wavefunction see lecture A3 by M. Betzinger.

4.2 Approximations to the Hamiltonian

Instead of trying to find approximations for the many-electron wavefunction alternatively one could find approximations to the Hamiltonian. In order to approximate the Hamiltonian appropriately one needs to have an idea, prior to the calculation, of the physical effects one is looking for. The idea is then to keep the necessary ingredients of the Hamiltonian to describe these effects while reducing the complexity of the Hamiltonian to a level where one can solve the Schrödinger equation. Frequently, the following two basic approximations are made: (i) only a

limited number of electrons are considered, i.e. one tries to identify the few degrees of freedom of the system which describe the relevant phenomena, and (ii) the two-particle electron-electron interaction is assumed to be a (semi-)local interaction which only acts between electrons in orbitals at the same lattice site or at close-by lattice sites, i.e. instead of $V_{ee}(\mathbf{r}, \mathbf{r}')$ one considers a U_{ij}, where i and j label such orbitals. Such model Hamiltonians are usually written in second quantization (see [12] and lecture B1 by E. Koch for an introduction into the formalism).
For example, in order to describe a single impurity atom in a metal one reduces the impurity atom to a single energy level which can be filled with maximally two electrons described by the Hamiltonian

$$\hat{H}_{\text{impurity}} = \sum_{\sigma} (\epsilon_d - \mu)\, \hat{c}^{\dagger}_{d\sigma} \hat{c}_{d\sigma} + U \hat{n}_{d\uparrow} \hat{n}_{d\downarrow}, \tag{46}$$

where $\hat{c}^{\dagger}$ and $\hat{c}$ are the usual electron creation and annihilation operators and $\hat{n}_{d\sigma}$ denotes the particle number operator for particles with spin σ on the impurity level d. In case the impurity level is occupied with two electrons, one of each spin, an interaction energy of size U is added. The electrons in the metal are described by momentum eigenstates, i.e. plane-waves with momentum $\mathbf{k}$, and are treated as non-interacting. The corresponding Hamiltonian, therefore, reads

$$\hat{H}_{\text{metal}} = \sum_{\mathbf{k}\sigma} (\epsilon_{\mathbf{k}} - \mu)\, \hat{c}^{\dagger}_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}\sigma}. \tag{47}$$

The impurity level and the metal are then coupled via a term that allows the electrons to hop from a level in the metal to the impurity level and vice versa with a probability $t_{\mathbf{k}}$ and $t^{*}_{\mathbf{k}}$, respectively. The coupling Hamiltonian is given by

$$\hat{H}_{\text{coupling}} = \sum_{\mathbf{k}\sigma} t_{\mathbf{k}} \hat{c}^{\dagger}_{d\sigma} \hat{c}_{\mathbf{k}\sigma} + t^{*}_{\mathbf{k}} \hat{c}^{\dagger}_{\mathbf{k}\sigma} \hat{c}_{d\sigma}. \tag{48}$$

Depending on the parameters $\epsilon_{d\sigma}$, $\epsilon_{\mathbf{k}\sigma}$, U, and $t_{\mathbf{k}}$ one finds the impurity level to be occupied with one or two electrons leading to a magnetic or non-magnetic impurity, respectively.
When describing magnetic systems with delocalized electrons one can use the Hubbard model. Here, the interaction is treated as an purely local interaction which results in a constant interaction in momentum space. The delocalized electrons are treated in the basis of plane-waves with momentum $\mathbf{k}$ such that the Hamiltonian for the system reads as

$$\hat{H}_{\text{Hubbard}} = \sum_{\mathbf{k}\sigma} \zeta_{\mathbf{k}} \hat{c}^{\dagger}_{\mathbf{k}\sigma} \hat{c}_{\mathbf{k}\sigma} + \frac{U}{2V} \sum_{\mathbf{k},\mathbf{k}',\mathbf{q}} \sum_{\sigma\sigma'} \hat{c}^{\dagger}_{\mathbf{k}+\mathbf{q}\sigma} \hat{c}^{\dagger}_{\mathbf{k}'-\mathbf{q}\sigma'} \hat{c}_{\mathbf{k}'\sigma'} \hat{c}_{\mathbf{k}\sigma}. \tag{49}$$

For a more detailed analysis of model Hamiltonians see lecture B1 by E. Koch.

4.3 Effective single-particle methods

If the Coulomb interaction was a small perturbation to the system or short-ranged and hence only affecting a small part of the system, we might be able to use perturbation theory to solve the interacting problem. Unfortunately, the Coulomb interaction is long-ranged and in many cases not a small effect. So the question arises if the single-particle description we discussed in some detail in Sec. 3 has any relevance for an realistic system in which the electrons interact. We already discussed that the electronic properties of a system of non-interacting electrons are predominantly governed by those electrons that occupy states close to the Fermi energy. Hence,

if the Coulomb interaction does not modify the properties of those electrons drastically, we stand a chance of still using the concepts of non-interacting electrons with some modifications. It was discovered by Landau [13, 14, 15] that one can indeed treat the electrons in many situations via the use of non-interacting particles in what became known as Fermi liquid theory. A crucial ingredient in this theory is the fermionic nature of the electrons as illustrated by the following example. Consider a system with all states up to the Fermi energy filled, a situation that happens hat zero temperature, and one additional electron occupying a state above the Fermi energy, i.e. $\epsilon_1 > \epsilon_F$. Now the electron-electron interaction we introduce scattering between the electrons, i.e. in other words the single particle description in which the electron is in some well-defined state will break down. For a scattering event to occur the second electron in the scattering process necessarily has an energy $\epsilon_2 \leq \epsilon_F$. Since the electrons need to scatter into empty states we obtain for the two energies after the scattering event: $\epsilon_3 \geq \epsilon_F$ and $\epsilon_4 \geq \epsilon_F$. Due to the conservation of energy we also have to satisfy $\epsilon_1+\epsilon_2 = \epsilon_3+\epsilon_4$. If the energy ϵ_1 we started the discussion with was exactly the Fermi energy, the only way to satisfy all these constraints would be $\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon_4$. In other words, the event would have zero weight in phase space and therefore not occur at all. For a particle with energy above the Fermi energy the probability of the event increases with increasing distance from the Fermi energy. All those effects that are determined by electrons around the Fermi energy the phase space remains reasonably small. This phase space argument relies only on the fact that fermions have to obey the Pauli exclusion principle and as such remains valid independent of the strength of the interaction. Hence, despite the fact that the Coulomb interaction is not a small perturbation the resulting disturbance of the electronic properties can still be small.

Generalizing this discussion to energies close to the Fermi-level, one finds a situation as sketched in Fig. 12. If the first electron changes its energy by $\Delta\epsilon = \epsilon_1 - \epsilon_3$ the phase-space of possible events is restricted by the condition $\Delta\epsilon < \epsilon_1 - \epsilon_F$ (cf. gray area in Fig. 12). The energy of the second electron then can be from an energy interval $[\epsilon_F - \Delta\epsilon, \epsilon_F]$ (light green area in Fig. 12). This fixes the whole scattering process and thus the total phase-space of the process will be $\propto (\epsilon_1 - \epsilon_F)^2$. This discussion indicates that the scattering rate (or inverse lifetime) of an electron with energy ϵ close to the Fermi-level will be proportional to $(\epsilon - \epsilon_F)^2$ and hence indicates again that the description of the system in terms of single particle levels actually is valid for these relevant electrons.

This discussion of scattering we presented so far of course assumes that the assignment of single electron states enables a valid description of the system. So in some sense it is a circular argument: if the single particle picture holds, the electron-electron interaction should not destroy the single particle picture. In fact, Landau put the argument on a more solid basis than this by explicitly studying the conditions under which one can obtain the interaction system of electrons (or more general fermions) starting from a non-interacting, so-called Fermi gas. These subtle arguments are beyond the scope of this simple introduction and we only want to point out that our very handwaving arguments actually can be put on firmer grounds.

In the process of converting a non-interacting (Fermi gas) system to an interacting (Fermi-liquid) description one also usually does not speak about single electrons anymore but introduces the idea of quasi-particle that have a finite lifetime. The quasi-particles can have different properties like a different mass (cf. our discussion of the effective mass in Sec. 3.2) but are still charged Fermions. The key point is the much reduced interaction between these quasi-particles. A description of the many electron state is then actually a picture in which weakly-interacting quasi-particles are approximated by a system of non-interacting particles. Before discussing such approximations we also should point out that there are effects in many-

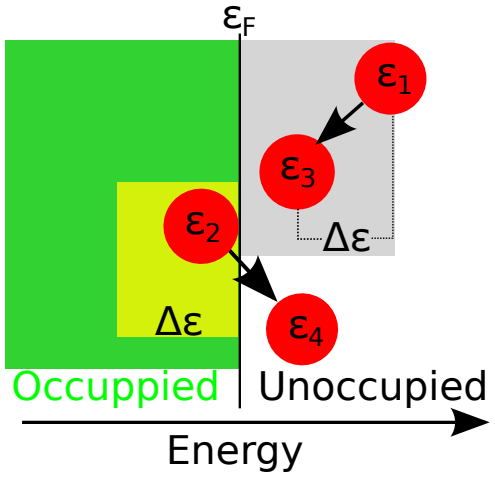


Fig. 12: *Scattering of two electrons. The first electron scatters from ϵ_1 to ϵ_3 and transfers energy $\Delta\epsilon$ to the second electron which scatters from ϵ_2 to ϵ_4.*

electron systems that can not be appropriately described by quasi-particles and that thus the Fermi-liquid description is limited. Consequently, such features are often associated with terms like 'non-Fermi-liquid behavior', 'non-quasi-particle physics', 'bad metal behavior'.

There exist several methods for constructing a system of effective single-particle equations that lead to an approximate solution of the many-body problem. Most successful of these approaches is density functional theory in which the effective non-interacting system is chosen such that it has the same ground-state density as the interacting many-body problem. It constructs an effective potential in which the single-particles move and which includes the effect of the electron-electron interaction. In this sense it is a "mean-field" approach, as the effect of all electrons on the particular state is treated by the potential or the field created. It can be shown that both the external potential and the effective single-particle potential are uniquely defined by the ground-state density [16]. As a consequence, all observables are functionals of the ground-state density which can be determined by minimizing the total energy. While large parts of the energy functional are known there is a small part, the so-called exchange-correlation energy, which needs to be approximated (For details see the manuscript A2 by M. Lezaic). Large parts of this Spring-school will actually use the picture provided by density functional theory and additional approximations and extensions.

Starting from a system of non-interacting quasi-particles one then can try to improve the description in a perturbative manner. This can be expected to be a reasonable approach due to the fact that the quasi-particles actually interact weakly. However, the calculation of the modified interaction itself is part of the complexity of such an approach. One usually employs Green-function methods to be able to deal effectively with the finite-lifetime of the quasi-particles and their modified interaction. Such an approach is discussed in the contribution A4 of Ch. Friedrich.

The methods introduced in this section will be discussed in more detail during other lectures and in the corresponding lecture notes. Due to the large variety of methods available, the selection presented here is far from complete. One method that is not addressed at all during the Spring school employs the one-body density matrix

$$\gamma(\mathbf{x},\mathbf{x}') = N \int d\mathbf{x}_2 \ldots \int d\mathbf{x}_N \Psi^*(\mathbf{x}' \ldots \mathbf{x}_N)\Psi(\mathbf{x} \ldots \mathbf{x}_N) \tag{50}$$

as its basic quantity building a functional theory known as reduced density matrix functional theory (RDMFT) on top of this variable [17, 18]. Written in its spectral representation the one-body density matrix reads as

$$\gamma(\mathbf{x},\mathbf{x}') = \sum_{j=1}^{\infty} n_j \varphi_j^*(\mathbf{x}')\varphi_j(\mathbf{x}). \tag{51}$$

with the eigenfunctions φ_j being known as natural orbitals and the eigenvalues n_j as their occupation numbers, the one-body density matrix retains the information whether the corresponding wavefunction Ψ is a Slater determinant or not. For a single Slater determinant the occupation numbers n_j only obtain two values, zero or one, leading to a so-called idempotent density matrix which satisfies

$$\int d\mathbf{y}\gamma(\mathbf{x},\mathbf{y})\gamma(\mathbf{y},\mathbf{x}') = \gamma(\mathbf{x},\mathbf{x}'). \tag{52}$$

Those orbitals φ_j which are occupied span the same space as the single-particle orbitals which compose the Slater determinant. The situation changes drastically when the wavefunction is comprised of more than one Slater determinant because the occupation numbers become fractional and might have any value between one and zero with their sum being identical to the particle number N.

In a sense RDMFT has its place somewhere in between single-particle approximations and wavefunction based methods. With the natural orbitals one still uses single-particle orbitals, however, the occupations are allowed to be fractional going clearly beyond the single Slater determinant approximation for the many-body wavefunction.

References

[1] N.W. Ashcroft, N.D. Mermin, *Solid State Physics* (Saunders College Publishing 1976).

[2] G.D. Mahan, *Many-particle physics* (Physics of Liquids and Solids, Springer 2000).

[3] A. Szabo, N.S. Ostlund, *Modern Quantum Chemistry*, (Dover Publications, 1996)

[4] E.K.U. Gross, E. Runge, O. Heinonen, *Many-Particle Theory*, (IOP Publishing, Bristol, Philadelphia and New York 1991).

[5] For an overview of all lattice structures and Bravais lattices, we refer to any introductory book on solid state theory or to wikipedia (http://en.wikipedia.org/wiki/Crystal_system)

[6] V. Heine, Surf. Sci. **2**, 1 (1964)

[7] D. Wortmann, H. Ishida, and S. Blügel, Phys. Rev. B **65**, 165103 (2002)

[8] N. Marzari, and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[9] W. Kohn, Phys. Rev. **115**, 809 (1959).

[10] N. Marzari, A. Mostofi, J. Yates, I. Souza, D. Vanderbilt, Rev. Mod. Phys. **84**, 1419 (2012).

[11] The Wannier90 code and its documentation is a good starting point for further information on details regarding the actual construction of Wannier functions (http://www.wannier.org).

[12] A.L. Fetter, J.D. Walecka, *Quantum Theory of Many-Particle Systems*, (Dover Publications, 2003).

[13] L.D. Landau, ZhETF (USSR) **30** (1956)

[14] G. Baym, and C. Pethick, *Landau Fermi-Liquid Theory* , (Wiley-VCH, 2004).

[15] L.D. Landau, E.M. Lifshitz, *Statistical Physics Part 2* (Butterworth-Heinemann, Oxford 1980)

[16] P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964).

[17] T.L. Gilbert, Phys. Rev. B **12**, 2111 (1975).

[18] A.M.K. Müller, Phys. Rev. A **105**, 446 (1984).

A 2 Introduction to Density Functional Theory[1]

Marjana Ležaić and Yuriy Mokrousov
Peter Grünberg Institut (PGI-1) and Institute for Advanced Simulation (IAS-1)
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Calculating the electronic properties of solids is not a trivial task, given the fact that the atomic nuclei and the electrons constitute a complex many-body problem. Therefore, all theories that deal with these calculations start by adopting the Born-Oppenheimer (adiabatic) approximation, which simply neglects the movement of the atomic nuclei while evaluating the electronic properties. This is a pretty good assumption, since the electrons are much lighter than the nuclei and thus move much faster. One can now focus solely on the electrons, which in itself is a formidable problem. The electrons interact with the positive atomic nuclei and with each other via Coulomb forces. Although the former interaction is by no means simple it can be treated, whereas the latter interaction is impossible to calculate and one must resort to approximations. Attempts to estimate the electron-electron interaction in solids and calculate the electronic dispersion or the total energy of different systems date back to the days of the Thomas-Fermi model [1, 2], the Hartree approximation [3] and to the X-α method of Slater [4]. The extension of these ideas which brought a revolution in the parameter-free *ab-initio* description of complex electronic structure is known as Density Functional Theory (DFT) and was established by Hohenberg and Kohn [5] and Kohn and Sham [6]. This has made it possible to calculate the total energy of solids, using the electron density, $n(\mathbf{r})$, as the key variable (for magnetic systems one has also to consider the magnetization density, $\mathbf{m}(\mathbf{r})$). In this chapter the Density Functional Theory will be described in more detail.

2 Electronic density instead of the wavefunction

The basic ideas of DFT will be presented here by considering a non-magnetic system with spin degeneracy first, while the spin polarized case will be discussed subsequently.

2.1 The Hohenberg-Kohn theorem

DFT is based on two theorems that provide the basis for the substitution of the many-body wavefunction with an electronic density in a description of a quantum-mechanical system.

Theorem 1 *For a given external potential v, the total energy of a system is a unique functional of the ground state electron density.*

To prove this we consider a Hamiltonian, $\hat{H} = \hat{T} + \hat{V} + \hat{V}_{ee}$, where $\hat{T}$ represents the kinetic energy of the system, $\hat{V}$ the interaction of the electrons with an external potential (including the potential coming from the atomic nuclei in the solid) and $\hat{V}_{ee}$ the electron-electron interaction. The solution to this Hamiltonian results in a ground state many body wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N)$ (for N electrons), and we have

$$\hat{H}\Psi = E_0\Psi. \tag{1}$$

The electron density can be calculated from

$$n(\mathbf{r}) = \left\langle \Psi \left| \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i) \right| \Psi \right\rangle, \tag{2}$$

and the interaction $\hat{V}$ is written as $V = \int n(\mathbf{r}) v_{ext}(\mathbf{r}) d^3r$, where $v_{ext}(\mathbf{r})$ is the external potential. What follows is the proof that two different external potentials $v_{ext}(\mathbf{r})$ and $v'_{ext}(\mathbf{r})$ must give rise to different ground state electron densities. For a system with potential $v'_{ext}(\mathbf{r})$ we have

$$\hat{H}'\Psi' = E'_0\Psi'. \tag{3}$$

From the variational principle it follows that

$$E_0 = \langle \Psi \mid \hat{H} \mid \Psi \rangle < \langle \Psi' \mid \hat{H} \mid \Psi' \rangle. \tag{4}$$

By adding and subtracting $v'_{ext}(r)$ on the rhs. of Eqn. 4 we obtain

$$\begin{aligned} \langle \Psi' \mid \hat{H} \mid \Psi' \rangle &= \langle \Psi' \mid \hat{H}' + \hat{V} - \hat{V}' \mid \Psi' \rangle \\ &= E'_0 + \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r}))d^3r. \end{aligned} \tag{5}$$

Combining the expressions in Eqns. 4 and 6 gives

$$E_0 < E'_0 + \int n'(\mathbf{r})(v(\mathbf{r}) - v'(\mathbf{r}))d^3r. \tag{6}$$

A similar argument, starting from the expression

$$E'_0 = \langle \Psi' \mid \hat{H}' \mid \Psi' \rangle < \langle \Psi \mid \hat{H}' \mid \Psi \rangle, \tag{7}$$

results in

$$E'_0 < E_0 + \int n(\mathbf{r})(v'_{ext}(\mathbf{r}) - v_{ext}(\mathbf{r}))d^3r. \tag{8}$$

Adding Eqn. 6 and 8 and assuming $n'(\mathbf{r}) = n(\mathbf{r})$ one obtains

$$E_0 + E'_0 < E'_0 + E_0, \tag{9}$$

which is clearly wrong. Hence $n'(\mathbf{r}) \neq n(\mathbf{r})$ and we conclude that two different potentials, $v_{ext}(\mathbf{r})$ and $v'_{ext}(\mathbf{r})$ give rise to different densities $n(\mathbf{r})$ and $n'(\mathbf{r})$. Therefore, knowledge of the electron density, $n(\mathbf{r})$, implies that it was calculated from a Hamiltonian with a specified external potential $v_{ext}(\mathbf{r})$. As the kinetic energy, $\hat{T}$, and electron-electron interactions, $\hat{V}_{ee}$, are known and specified one concludes that knowledge of the ground state electron density determines the entire Hamiltonian and hence the ground state energy, which proves Theorem 1. One can thus express a functional relationship between the ground state energy and the corresponding electron density as

$$E[n(\mathbf{r})] = T[n(\mathbf{r})] + V[n(\mathbf{r})] + V_{ee}[n(\mathbf{r})]. \tag{10}$$

The second important theorem of DFT is

Theorem 2 *The exact ground state density minimizes the energy functional* $E[n(\mathbf{r})]$.

To prove Theorem 2 one starts from Theorem 1 and for a given external potential $v_0(\mathbf{r})$ writes

$$E_{v_0}[n(\mathbf{r})] = \langle \Psi[n(\mathbf{r})] \mid \hat{T} + \hat{V}_{ee} + \hat{V}_0 \mid \Psi[n(\mathbf{r})] \rangle, \tag{11}$$

where the subscript v_0 indicates that this is the energy functional for a system with external potential $v_{ext} = v_0(\mathbf{r})$. Since the ground state density specifies the Hamiltonian, it also specifies the wave function (of the ground state and of excited states) and hence the notation $\Psi[n(\mathbf{r})]$. If the ground state electron density is denoted by $n_0(\mathbf{r})$, the ground state can be expressed as $\Psi[n_0(\mathbf{r})]$. From the variational principle one again obtains

$$\langle \Psi[n_0(\mathbf{r})] \mid \hat{T} + \hat{V}_{ee} + \hat{V}_0 \mid \Psi[n_0(\mathbf{r})] \rangle < \langle \Psi[n(\mathbf{r})] \mid \hat{T} + \hat{V}_{ee} + \hat{V}_0 \mid \Psi[n(\mathbf{r})] \rangle, \tag{12}$$

which can also be expressed as

$$E_{v_0}[n_0(\mathbf{r})] < E_{v_0}[n(\mathbf{r})], \tag{13}$$

i.e., the ground state density minimizes the energy functional $E[n(\mathbf{r})]$, which is what Theorem 2 states. If one would now have an explicit form for $E[n(\mathbf{r})]$ they could go ahead and minimize it with respect to the electron density and in this way calculate the ground state energy. Unfortunately, due to the complexity provided by the electron-electron interactions, approximations are necessary to obtain an explicit expression for $E[n(\mathbf{r})]$.

2.2 DFT for spin polarized systems

The reasoning applied to the non-magnetic systems can be extended to the spin polarized ones, and it can be shown that the ground state energy is a unique functional of the electron and magnetization density, $n(\mathbf{r})$ and $\mathbf{m}(\mathbf{r})$, with

$$\mathbf{m}(\mathbf{r}) = -\mu_B \left\langle \Psi \left| \boldsymbol{\sigma} \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i) \right| \Psi \right\rangle. \tag{14}$$

Here, μ_B is the Bohr magneton and $\boldsymbol{\sigma}$ is the vector of Pauli matrices. The proof of this is quite similar to the proof outlined above. One starts by modifying the Hamiltonian to include an external magnetic field, $\mathbf{B}_{ext}(\mathbf{r})$, so that the Hamiltonian becomes $\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{U}$, where

$$U = \int \left[v_{ext}(\mathbf{r}) n(\mathbf{r}) - \mathbf{B}_{ext}(\mathbf{r}) \cdot \mathbf{m}(\mathbf{r}) \right] d^3r. \tag{15}$$

Based on the variational principle, similar to the discussion around Eqns. 4-6, one arrives at

$$E_0 < E_0' + \int n'(\mathbf{r})[v_{ext}(\mathbf{r}) - v'_{ext}(\mathbf{r})]d^3r - \int \mathbf{m}'(\mathbf{r})[\mathbf{B}_{ext}(\mathbf{r}) - \mathbf{B}'_{ext}(\mathbf{r})]d^3r \tag{16}$$

and

$$E_0' < E_0 + \int n(\mathbf{r})[v'_{ext}(\mathbf{r}) - v_{ext}(\mathbf{r}))d^3r - \int \mathbf{m}(\mathbf{r})(\mathbf{B}'_{ext}(\mathbf{r}) - \mathbf{B}_{ext}(\mathbf{r})]d^3r. \tag{17}$$

Assuming that $n(\mathbf{r}) = n'(\mathbf{r})$ and $\mathbf{m}(\mathbf{r}) = \mathbf{m}'(\mathbf{r})$, and adding Eqns. 16 and 17 the same absurd result as in the discussion of spin degenerate systems, i.e. Eqn. 9 follows, and one must draw the conclusion that $n(\mathbf{r}) \neq n'(\mathbf{r})$ and $\mathbf{m}(\mathbf{r}) \neq \mathbf{m}'(\mathbf{r})$. Hence, for magnetic systems the ground state energy is a unique functional of the electron density and the magnetization density.

2.3 Thomas-Fermi approximation

Within the Thomas-Fermi (TF) approximation the electrons are considered to be independent and the electron-electron interaction energy between them comes only from the electrostatics:

$$E_{\rm es}[n] = \frac{e^2}{2} \int d\mathbf{r} \int d\mathbf{r}' \, \frac{n(\mathbf{r})\, n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{18}$$

Also, the kinetic energy is assumed to be of the form:

$$T[n] = \int t[n(\mathbf{r})]\, d\mathbf{r}, \tag{19}$$

where $t[n]$ is the kinetic energy of a system of independent electrons with density n, which is given by:

$$t[n] = \frac{2}{(2\pi)^3} \int_{k<k_F} \frac{\hbar k^2}{2m}\, d\mathbf{k}. \tag{20}$$

The density can be related to the magnitude of the Fermi Bloch vector k_F via:

$$\frac{2 \cdot \frac{4\pi}{3} \cdot k_F^3}{(2\pi)^3} = n. \tag{21}$$

This gives for the kinetic energy

$$T[n] = C_k \int n^{5/3}(\mathbf{r})\, d\mathbf{r}, \tag{22}$$

with $C_k = 3\hbar^2(3\pi^2)^{2/3}/10m$. We are ready now to minimize the total energy functional $E[n]$ under the condition that the number of electrons is constant:

$$\int n(\mathbf{r})\, d\mathbf{r} = N. \tag{23}$$

Applying the method of Lagrange multipliers we seek for a minimum of

$$E[n] + \lambda N = T[n] + E_{\rm es}[n] + \int n(\mathbf{r})[v(\mathbf{r}) + \lambda]\, d\mathbf{r}, \tag{24}$$

which results in the following Euler equation (Thomas-Fermi equation):

$$\frac{5}{3} C_k n(\mathbf{r})^{2/3} + e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' + v(\mathbf{r}) + \lambda = 0. \tag{25}$$

The Thomas-Fermi approximation, although it was used frequently in the past, has severe drawbacks. Eventhough it provides a comparatively good approximation to the charge density in some cases, it can be shown that the TF charge density is infinite at the nuclei, and it decays as r^{-6} away from them, and not exponentially. Moreover, within the TF approximation the binding of molecules or solids cannot be achieved, and the periodicity in the properties of solids with respect to the nuclear number is not reproduced. The TF approach also never leads to ferromagnetism [7]. Despite all the negative sides of this approximation, it was of considerable interest in the past since it presents an exact solution in the limit of infinite nuclear charge.

3 The Kohn-Sham equations

How do we use DFT now to describe a solid? In the approach Kohn and Sham took, one starts with a simple non-interacting electron system where the part of the Hamiltonian describing electron-electron interactions, V_{ee}, is absent. In this case the electrons which move in the field of an external potential which, for reasons that will be obvious below, will be called "effective" potential v_s, are solutions to a one-electron Schrödinger equation,

$$\left(-\frac{\nabla^2}{2} + v_s\right)\varphi_i = \epsilon_i \varphi_i. \tag{26}$$

There is an infinite number of solutions to this equation that are labeled by the subscript i. From Eqn. 26 one can calculate an electron density from the lowest lying one-particle (op) states. If there are N electron states which are solutions to Eqn. 26 one simply calculates the one-particle (a label introduced to show that there are no electron-electron interactions considered) electron density from

$$n_{op}(\mathbf{r}) = \sum_{i=1}^{N/2} 2 \mid \varphi_i(\mathbf{r}) \mid^2, \tag{27}$$

where the factor 2 comes from spin degeneracy. In this case the energy functional which describes the total energy of the N electrons may be written as,

$$\begin{aligned} E_{op}[n_{op}(\mathbf{r})] &\equiv T_{op}[n_{op}] + v_s[n_{op}] \\ &= \sum_{i=1}^{N/2} \left\langle \varphi_i(\mathbf{r}) \left| \frac{-\nabla^2}{2} \right| \varphi_i(\mathbf{r}) \right\rangle + \int n_{op}(\mathbf{r}) v_s(\mathbf{r})\, d\mathbf{r}, \end{aligned} \tag{28}$$

and the electron density which minimizes this functional is obtained from the requirement that the energy functional is stationary for small variations of the electron density around the ground state density. This can be written as

$$0 = \delta E_{op} = E_{op}[n_{op}(\mathbf{r}) + \delta n_{op}(\mathbf{r})] - E_{op}[n_{op}(\mathbf{r})], \tag{29}$$

which may also be written as

$$0 = \delta T_{op}[n_{op}] + \int \delta n(\mathbf{r}) v_s(\mathbf{r})\, d\mathbf{r}. \tag{30}$$

Carrying out the minimization in Eqn. 30 leads to Eqn. 26. Thus, the independent particles which are the solution to Eqn. 26 give rise to a density which minimizes the total energy expression of independent particles in Eqn. 29. The reason for introducing Eqns. 26 to 30 is motivated by the fact that they can be solved, at least approximately, to within a desired accuracy. More important, however, is the fact that they can, via the Kohn-Sham approach, be used to actually calculate the ground state energy of a "real" interacting electron system. The basic principle of the Kohn-Sham approach is now to assume that one can find an effective potential, v_s, so that $n_{op}(\mathbf{r}) = n(\mathbf{r})$ where $n(\mathbf{r})$ is the electron density of the fully interacting system. Since we know that the total energy of a system is uniquely determined by the electron density, it seems to be an efficient route to obtain the correct electron density from a one-electron like problem.

The question now is how to determine v_s, so that $n_{op}(\mathbf{r})$ becomes equal to $n(\mathbf{r})$. To do this one rewrites the energy functional in Eqn. 10 as

$$\begin{aligned} E[n(\mathbf{r})] &= T_{op}[n(\mathbf{r})] + \int n(\mathbf{r}) v_{ext}(\mathbf{r})\, d\mathbf{r} \\ &+ \frac{e^2}{2} \int\int \frac{n(\mathbf{r}) \cdot n(\mathbf{r}')}{\mid \mathbf{r} - \mathbf{r}' \mid} d\mathbf{r}\, d\mathbf{r}' + E_{xc}[n(\mathbf{r})] \end{aligned} \tag{31}$$

(since it is required that $n_{op}(\mathbf{r})$ should be equal to $n(\mathbf{r})$ in the expression above, for simplicity, the subscript *op* on the right hand side was skipped). In Eqn. 31 the one-particle kinetic energy functional instead of the true kinetic energy functional of Eqn. 10 was introduced. The electrostatic electron-electron interactions are contained in the so-called Hartree term (third term on the right-hand side). This term we saw already in the Thomas-Fermi approximation Eq. 18. The remaining electron-electron interactions and the corrections to the kinetic energy are gathered in the exchange-correlation term E_{xc}. Hence in order to make Eqn. 31 equal to Eqn. 10 one must introduce a correcting term, and this is what the exchange and correlation energy, $E_{xc}[n(\mathbf{r})]$, does. Since the first three terms on the right hand side of Eqn. 31 are possible to calculate numerically, in this way the problem of the complexity of the fully interacting system is mapped into the problem of finding the exchange and correlation functional. Using the ansatz for the total energy we can rewrite Eq. 29 as follows:

$$0 = \delta T_{op}[n] + \int d\mathbf{r}\, \delta n(\mathbf{r}) \left[v_{ext}(\mathbf{r}) + \int e^2 \frac{n(\mathbf{r}')}{\mid \mathbf{r} - \mathbf{r}' \mid} d\mathbf{r}' + \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})} \right], \tag{32}$$

from which the v_s can be readily identified:

$$v_s(\mathbf{r}) = v_{ext}(\mathbf{r}) + \int e^2 \frac{n(\mathbf{r}')}{\mid \mathbf{r} - \mathbf{r}' \mid} d\mathbf{r}' + \frac{\delta E_{xc}[n(\mathbf{r})]}{\delta n(\mathbf{r})}. \tag{33}$$

So far the exact exchange and correlation functional which would make Eqn. 31 hold for all densities and all systems, was not found. However, for a uniform electron gas one can calculate $E_{xc}[n(\mathbf{r})]$ for all values of the electron density and thus parametrized forms of $E_{xc}[n(\mathbf{r})]$ as a function of $n(\mathbf{r})$ can be built. Strictly speaking, this can be done in the high electron density limit [8] and in the low electron density limit [9]. Interpolation between these two limits gave rise to parametrized forms of the exchange and correlation functional of a uniform electron gas for all values of the density [9]. However, this interpolation in modern electronic structure calculations replaced by approaches which are based on quantum Monte-Carlo simulations for the intermediate values of the electron gas [10].

3.1 The local density approximation (LDA)

The local density approximation (LDA) assumes that the parametrizations of the exchange-correlation energy worked out for the case of uniform electron gas work even in cases when the electron gas is not uniform, but varies in space, as it does in a solid, surface or interface. In the frame of this approximation, the expression

$$E_{xc}[n(\mathbf{r})] = \int \epsilon_{xc}\left[n(\mathbf{r})\right] n(\mathbf{r})\, d\mathbf{r}, \tag{34}$$

for the exchange-correlation energy is introduced, where $\epsilon_{xc}[n(\mathbf{r})]$ is named the exchange-correlation energy density. In a parametrized form its dependence on $n(\mathbf{r})$ is relatively simple

and may, for example, be found in Ref. [11]. Armed with an (approximate) expression for the ground state energy functional and in analogy with Eqns. 29 and 30 the ground state density can be determined from this functional by requiring that the functional (Eqn. 31) is stationary for small variations of the electron density around the ground state density. As a result, an expression which is quite similar to Eqn. 33 follows:

$$v_s(\mathbf{r}) = v_{ext}(\mathbf{r}) + \int e^2 \frac{n(\mathbf{r}')}{\mid \mathbf{r} - \mathbf{r}' \mid} d\mathbf{r}' + \mu_{xc}(n(\mathbf{r})), \tag{35}$$

where

$$\mu_{xc}(n(\mathbf{r})) = \frac{\delta(E_{xc}[n(\mathbf{r})])}{\delta n(\mathbf{r})} = \epsilon_{xc}[n(\mathbf{r})] + n(\mathbf{r}) \frac{\partial(\epsilon_{xc}[n(\mathbf{r})])}{\partial n(\mathbf{r})}. \tag{36}$$

Eqn. 26 should be solved now with the effective potential specified by Eqn. 35. Since the effective potential to be used in Eqn. 26 depends on the electron density, the property that should be calculated, one has to perform a self-consistent calculation where an initial electron density is more or less guessed and an effective potential is calculated from Eqn. 35. This potential is then used to solve Eqn. 26 and a new electron density is calculated from Eqn. 27, which is then put back into Eqn. 35. This procedure is repeated until convergence is obtained, i.e. until the density does not change appreciably with successive iterations. In practice, one mixes the electron density which is the output of Eqn. 27 with the electron density which is in input for that particular loop in the self-consistency iterational procedure before one takes this mixed density and puts it in Eqn. 35. The whole procedure of mixing is quite complex where many suggestions of how to achieve self consistency with as few iterations as possible have been suggested [12, 13]. Once a self consistent electron density has been found one can calculate the ground state energy of the Kohn-Sham (LDA) energy functional (via Eqn. 31) and hence one of the main goals in electronic structure calculations has been achieved.

3.2 Exchange-correlation energy

One of the key points of the Kohn-Sham approach is building a correspondence between the density of the real system and the density of as system of non-interacting particles in an effective potential, governed by equations which can be solved in principle exactly for systems of great complexity. In order to inspect this correspondence, we consider a two-particle interaction

$$\frac{\lambda}{|\mathbf{r} - \mathbf{r}'|}, \tag{37}$$

where we can tune by hand parameter λ from 0 = non-interacting case, to 1 = real system. We have to do this in the presence of another external potential V_λ, which keeps the density of our imaginary system constant. The Hamiltonian of our λ-system reads:

$$H^\lambda = -\frac{\nabla^2}{2} + v_{ext} + V_\lambda + V_{\mathrm{ee}}^\lambda, \tag{38}$$

where V_{ee}^λ is the λ-scaled Coulomb interaction of the form above. Remarkably, the exchange-correlation energy for $\lambda = 1$ can be expressed as [14]:

$$E_{\mathrm{xc}} = \frac{1}{2} \int d\mathbf{r}\, n(\mathbf{r}) \int d\mathbf{r}'\, \frac{n_{\mathrm{ex}}(\mathbf{r}, \mathbf{r}' - \mathbf{r})}{|\mathbf{r} - \mathbf{r}'|}, \tag{39}$$

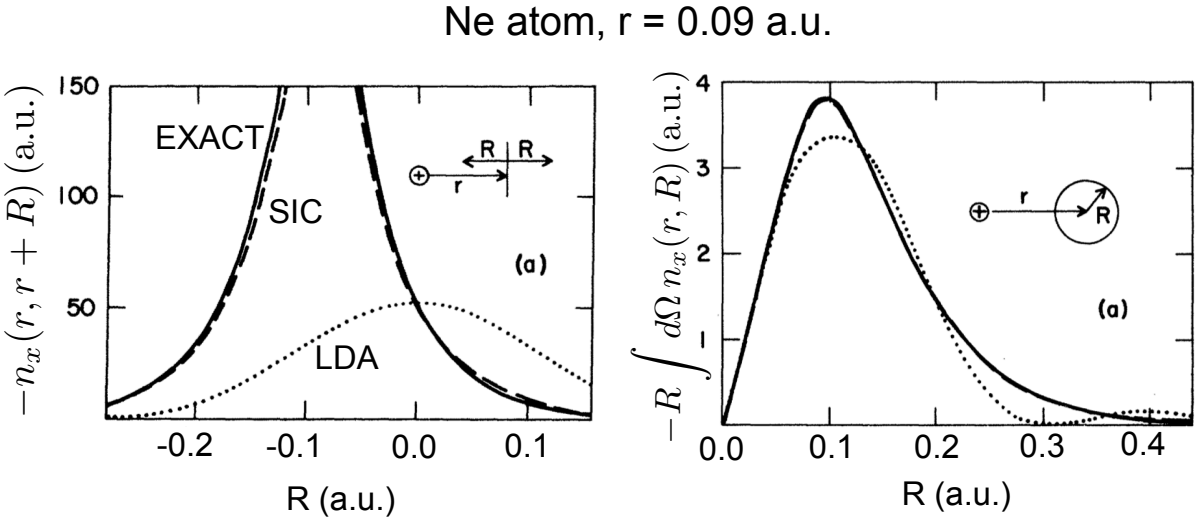


Fig. 1: *Left: exchange part $n_x(\mathbf{r}, \mathbf{r}+\mathbf{R})$ of the exchange-correlation hole defined in Eq. 40 for Ne atom at $|\mathbf{r}| = r = 0.09$ a.u. Note that $|\mathbf{R}| = R$. Right: corresponding spherical average of the exchange hole which enters Eq. 41. Solid line − exact result, dotted line − LDA result, dashed line − SIC-corrected LDA result. As can be seen, large difference between the exact and LDA exchange hole apparent in the left figure does not result in large difference between corresponding spherically-averaged values. Adapted from [15].*

where the so-called *exchange-correlation hole* is given by:

$$n_{\rm ex}(\mathbf{r}, \mathbf{r}' - \mathbf{r}) := n(\mathbf{r}') \int_0^1 (g(\mathbf{r}, \mathbf{r}', \lambda) - 1)\, d\lambda, \tag{40}$$

and it describes an effect of repulsion of electron at point $\mathbf{r}'$ felt due to the presence of electron at point $\mathbf{r}$. Function $g(\mathbf{r}, \mathbf{r}', \lambda)$ above is the pair correlation function of the system with density $n(\mathbf{r})$ and Coulomb interaction due to $V_{\rm ee}^{\lambda}$. It can be shown that $g(\mathbf{r}, \mathbf{r}', \lambda)$ tends to unity as the distance between $\mathbf{r}$ and $\mathbf{r}'$ increases. We can thus roughly say that the energy $E_{\rm xc}$ given by Eq. 39 stands for the short-range part of the Coulomb interaction, while the long-range part of it is reflected in the electrostatic part of the total energy, the expression for which we know exactly. An example of how an exchange-correlation hole looks like in real space is presented in Fig. 1 for a Neon atom. As can be seen in this figure, the LDA approximation for the exchange-correlation hole is in drastic difference to the exact result for its distribution. A reasonable question to ask is how can the local density approximation produce reasonable results given such a poor description of the exchange-correlation hole?

This question can be answered by writing $\mathbf{R} = \mathbf{r}' - \mathbf{r}$, and rewriting Eq. 39 in the following way:

$$E_{\rm ex} = \frac{1}{2} \int d\mathbf{r}\, n(\mathbf{r}) \int_0^{\infty} dR\, R^2\, \frac{1}{R} \int d\Omega\, n_{\rm xc}(\mathbf{r}, \mathbf{R}), \tag{41}$$

where $d\Omega$ stands for an infinitesimal solid angle. The latter expression points out to the fact, that, although the description of non-spherical components of $n_{\rm ex}(\mathbf{r}, \mathbf{R})$ can be very imprecise, we can still get a very good estimate for $E_{\rm ex}$, since it depends only on the spherical average of the exchange-correlation hole. See for example the comparison of the spherically averaged

exact and LDA exchange-correlation holes presented in Fig. 1. The difference between the left and right parts of this figure, as far as the comparison between the curves is concerned, is drastic. Finally, from the definition of the pair-correlation function one can deduce that the so-called *sum rule*, which should be satisfied for each $\mathbf{r}$:

$$\int d\mathbf{r}'\, n_{\rm xc}(\mathbf{r}, \mathbf{r}' - \mathbf{r}) = -1. \tag{42}$$

This means that for each $\mathbf{r}$ we can introduce the radius of the exchange-correlation hole:

$$\left\langle \frac{1}{R} \right\rangle_{\mathbf{r}} = -\int d\mathbf{R}\, \frac{n_{\rm xc}(\mathbf{r}, \mathbf{R})}{R}, \tag{43}$$

with the consequent expression for the exchange-correlation energy in terms of the $\mathbf{r}$-dependent radius:

$$E_{\rm ex} = \frac{1}{2} \int d\mathbf{r}\, n(\mathbf{r}) \left\langle \frac{1}{R} \right\rangle_{\mathbf{r}}. \tag{44}$$

This emphasizes again that $E_{\rm ex}$ depends very loosely on the exact distribution of the exchange-correlation hole in real space.

3.3 Kohn-Sham equations for spin-polarized systems

At the end of the previous section it was shown that for magnetic systems the ground state energy may be written as a unique functional of the electron density and of the magnetization density. An alternative way of expressing this is to state that there is an energy functional which depends both on the majority and the minority spin density (since $n(\mathbf{r}) = n^{\uparrow}(\mathbf{r}) + n^{\downarrow}(\mathbf{r})$ and $m(\mathbf{r}) = n^{\uparrow}(\mathbf{r}) - n^{\downarrow}(\mathbf{r})$[2] and the energy functional can be written as $E[n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})]$. Following arguments similar to the ones around Eqn. 31, one obtains a Kohn-Sham scheme for spin-polarized systems through

$$\begin{aligned} E[n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})] &= T_{op}[n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})] + \int n(\mathbf{r}) v_{ext}(\mathbf{r}) d^3r \\ &+ \frac{1}{2} \int\int e^2 \frac{n(\mathbf{r}) \cdot n(\mathbf{r}')}{\mid \mathbf{r} - \mathbf{r}' \mid} d^3r d^3r' + E_{xc}[n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})]. \end{aligned} \tag{45}$$

In a real solid the preference for occupying one spin channel (to some degree) more than the other is traditionally explained as due to the exchange interaction and the driving force for it is the electron-electron interaction in the Hamiltonian. Hence in the spin-polarized Kohn-Sham scheme this necessarily means that the exchange and correlation potential, which is supposed to absorb all complex electron-electron interactions, must depend both on the charge and the spin (magnetization) density. Turning again to studies on the uniform electron density is useful and parametrizations for $E_{xc}[n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})]$, as a function of $n^{\uparrow}(\mathbf{r})$ and $n^{\downarrow}(\mathbf{r})$, have been made. Proceeding analogously to the discussion around 26 and 27, one analyzes a one-particle Hamiltonian with spin up (down) effective potentials,

$$\left(-\frac{\nabla^2}{2} + v_s^{\uparrow(\downarrow)} \right) \varphi_i^{\uparrow(\downarrow)} = \epsilon_i^{\uparrow(\downarrow)} \varphi_i^{\uparrow(\downarrow)}, \tag{46}$$

[2]This approach, however, simplifies the situation somewhat since the magnetization density is a vector property with both magnitude and direction. In this analysis it is assumed that the magnetization is pointing only in one direction, the z-direction, of the system and only its magnitude is considered.

where the electron density for electrons with a given spin is obtained from

$$n_{op}^{\uparrow(\downarrow)}(\mathbf{r}) = \sum_{i=1} \mid \varphi_i^{\uparrow(\downarrow)}(\mathbf{r}) \mid^2 . \tag{47}$$

Repeating the discussion which led to Eqn. 35, with the only modification that one now requires the energy functional to be stationary with regard to both the spin up and the spin down density, leads to effective potentials which are different for the two spin directions due to differences in the exchange and correlation potential,

$$v_s^{\uparrow(\downarrow)}(\mathbf{r}) = v(\mathbf{r}) + \int e^2 \frac{n(\mathbf{r}')}{\mid \mathbf{r} - \mathbf{r}' \mid} d^3r' + \mu_{xc}^{\uparrow(\downarrow)}(n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})), \tag{48}$$

where the exchange-correlation potential is defined similar to Eqn. 36 as

$$\mu_{xc}^{\uparrow(\downarrow)}(n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})) = \frac{\delta E_{xc}[n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r})]}{\delta n^{\uparrow(\downarrow)}(\mathbf{r})}. \tag{49}$$

Hence, the simplest forms of spin polarized calculations treat the spin up and spin down electrons separately and for every iteration in the self consistent loop one solves a Kohn-Sham equation for both spin directions. The spin up and spin down densities are then calculated by occupying the N lowest (spin up or spin down) eigenvalues of the separate two Kohn-Sham equations. Since for a given $v_s^{\uparrow}(\mathbf{r})$ which may be different from $v_s^{\downarrow}(\mathbf{r})$ there may be more spin up states, $\epsilon_i^{\uparrow}$ than spin down states, $\epsilon_i^{\downarrow}$, which have an energy lower than the highest occupied state (the Fermi level, E_F) it is clear how spin polarization might occur. With a self consistent spin and magnetization density the magnetic moment is calculated as $\int m(\mathbf{r}) d^3r$ (in Bohr magneton units) and the total energy may be calculated from Eqn. 45.

3.4 The Generalized Gradient Approximation (GGA)

Within the LDA the exchange-correlation energy functional can be written in terms of the $\varepsilon_{xc}(n^{\uparrow}, n^{\downarrow})$, which is the exchange-correlation energy density per particle of an electron gas with uniform spin densities $n^{\uparrow}, n^{\downarrow}$ and $n = n^{\uparrow} + n^{\downarrow}$. This approximation is most appropriate when the spin densities vary slowly in space. However, this condition is not really satisfied in real atoms, molecules and solids, and we have attempted to explain why LDA often leads to a very good description of such systems by referring to the insensitivity of the exchange-correlation energy on the details of the exchange-correlation hole. Nevertheless, rigorously speaking, the next level of approximation which could also explicitly take into account the spatial variation of the density would be the approximation to the exchange-correlation functional, which is called the generalized gradient approximation (GGA), of the form:

$$E_{xc}^{GGA}[n^{\uparrow}, n^{\downarrow}] = \int f(n^{\uparrow}, n^{\downarrow}, \nabla n^{\uparrow}, \nabla n^{\downarrow})\, d\mathbf{r}. \tag{50}$$

We note that while different LDAs are simply different parametrizations, different GGAs are in fact different exchange-correlation functionals. A lot of research has been conducted in order to find expressions for f which would compare well with many other functionals and experiments over wide ranges of materials. One of the most successful and practically popular suggestions for f was made by Perdew, Burke and Erzernhof (PBE functional) [16].

3.5 Some practical aspects in solving the Kohn-Sham equations

3.5.1 Systems with translational symmetry

In the procedure of solving the Kohn-Sham equations in systems with translational invariance, several simplifications can be used due to this symmetry [17]. The discussion here will concern the bulk calculations, but the conclusions can also be applied to the lower dimensional systems, like thin films or nanowires, in the directions in which the translational symmetry is preserved. First of all one normally assumes in a bulk material that the potential which enters Eqn. 46 is periodic, i.e. $v_s^{\uparrow(\downarrow)}(\mathbf{r}) = v_s^{\uparrow(\downarrow)}(\mathbf{r}+\mathbf{R})$, where $\mathbf{R}$ is a translation vector (a Bravais lattice vector) of the solid. This periodic boundary condition leads to Bloch's theorem [17] which states that as an effect of the periodicity of the bulk material the one-electron wave function must obey the condition

$$\varphi_{i,\mathbf{k}}^{\uparrow(\downarrow)}(\mathbf{r}+\mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}}\varphi_{i,\mathbf{k}}^{\uparrow(\downarrow)}(\mathbf{r}), \tag{51}$$

where **k** is a vector of reciprocal space.[3] Due to the translation symmetry one has only to consider **k**-vectors which lie inside the first Brillouin zone[4] when looking for solutions to Eqns. 46-48 [17]. In addition one can solve the Eqns. 46-48 for each **k**-vector being separate and independent of the others. However, the dependence of the one-electron wave function on **k** makes the calculation of the one-electron density somewhat more complex since a sum over all possible **k**-vectors has to be included, and Eqn. 47 is in a crystal replaced by

$$n_{op}^{\uparrow(\downarrow)}(\mathbf{r}) = \sum_i \sum_{\mathbf{k}} \mid \varphi_{i,\mathbf{k}}^{\uparrow(\downarrow)}(\mathbf{r}) \mid^2 . \tag{52}$$

Similarly, in the calculation of the total energy, Eqn. 45, one needs to calculate the sum of Kohn-Sham eigenvalues ϵ_i (Eqn. 26),

$$E_{eig} = \sum_i \sum_{\mathbf{k}} \epsilon_{i,\mathbf{k}}, \tag{53}$$

since $T_{op} = E_{eig} - \int v_s(\mathbf{r}) n(\mathbf{r}) d^3r$. In principle, all **k**-vectors inside the first Brillouin zone (BZ) should be considered in the sums in Eqns. 52 and 53, but since this number is enormous one would like to replace the sum with an integral. However, if one does not have an analytic dependence of the Kohn-Sham eigenvalues on **k**, a way to approximate Eqn. 53 must be found. In order to do this it is useful to introduce the concept of density of states (DOS), which can be calculated from

$$D(E) = \sum_i \frac{1}{8\pi^3} \int_{BZ} \delta(E - \epsilon_{i\mathbf{k}}) d^3k. \tag{54}$$

With this definition of the DOS one can calculate the eigenvalue sum from

$$E_{eig} = \int_{-\infty}^{E_F} E D(E) dE. \tag{55}$$

The expressions 54 and 55 can be combined into

$$E_{eig} = \sum_i \int_{-\infty}^{\infty} E f(E) \frac{1}{8\pi^3} \int_{BZ} \delta(E - \epsilon_{i\mathbf{k}}) d^3k dE, \tag{56}$$

[3] Reciprocal space is spanned by the vectors $\mathbf{G}_i$, defined as $\mathbf{G}_i \cdot \mathbf{R}_j = 2\pi\delta_{ij}$, where V is the volume of the primitive cell of the Bravais lattice.

[4] Moreover, when the system under consideration additionally possesses a point group symmetry, only the irreducible wedge of the first Brillouin zone (determined by the symmetry) has to be considered.

where $f(E)$ is a step function which attains the value one below the Fermi energy and zero above. Finally, the Eq.56 is solved on a **k**-points mesh, distributed as to fulfill the symmetry of the space-group. The integral over the Brillouin zone can then be calculated as a weighted sum over the bands, i, and the discrete set of sampled **k**-points, $\mathbf{k}_j$, with weight functions, w_{ji}. Several methods can be used for interpolation of the eigenvalues between the **k**-points of the sample, like the linear tetrahedron method [18], the modified tetrahedron method [19], or the Gaussian broadening method [20].

3.5.2 Andersen's Force Theorem

It is quite often that one needs to calculate the difference in total energies between an initial and a slightly perturbed state. Andersen's force theorem claims that when the perturbation of the densities $n(\mathbf{r})$, $m(\mathbf{r})$ is small, the difference can be calculated as a difference in the sums of the eigenvalues of the two states (Eqn. 53). This can be shown as follows (for more details see [21]). The total energy can be written as

$$E = \int_{-\infty}^{E_F} ED(E)dE - E_r, \tag{57}$$

where $D(E)$ is the density of states and the first term on the right hand side (called *the band term*) stands for the eigenvalues sum, E_{eig} (Eqn. 55). The term E_r includes all the other contributions to the total energy and can be easily explicitly obtained from Eqn. 45 with $T_{op} = E_{eig} - \int v_s(\mathbf{r})n(\mathbf{r})d^3r$. The perturbation will cause some changes in the system and we assume these changes to be parametrized by a set of quantities $\{X_i\}$. The total energy is then a function of these parameters and the aim is to calculate

$$\delta E = \sum_i \frac{\delta E}{\delta X_i}\delta X_i. \tag{58}$$

Consider now a perturbation specified by a set of parameters $\{\Delta X_i\}$. The system can be led to self-consistency in two steps. First, the effective potential is held fixed and the Schrödinger equation is solved (for the new set of parameters $\{X_i + \Delta X_i\}$), giving rise to a new density and new energy eigenvalues. This variation we shall denote as δ_1. In the second step the potential is allowed to relax to self-consistency, with the same parameter set $\{X_i + \Delta X_i\}$. Denoting this variational step as δ_2, we can write the complete variation as $\delta = \delta_1 + \delta_2$. Then, the change in the band term in Eqn. 57 is

$$\delta E_{eig} = \delta_1 E_{eig} + \delta_2 E_{eig}, \tag{59}$$

and the change of the total energy is

$$\delta E = \delta_1 E_{eig} + \delta_2 E_{eig} - \delta E_r. \tag{60}$$

After writing out the term δE_r and applying the first-order perturbation theory for the term $\delta_2 E_{eig}$ [22], one finally obtains

$$\delta E = \delta_1 E_{eig} - \int_{\delta\Omega} n^2 \frac{\mathrm{d}\epsilon_{\mathrm{xc}}}{\mathrm{d}n}\delta\mathbf{S}\cdot\mathrm{d}\mathbf{S}, \tag{61}$$

where $\mathrm{d}\mathbf{S}$ is the surface element and $\delta\mathbf{S}$ describes the change in the volume Ω in the sense that a point $\mathbf{S}$ is taken to $\mathbf{S} + \delta\mathbf{S}$ by the perturbation.

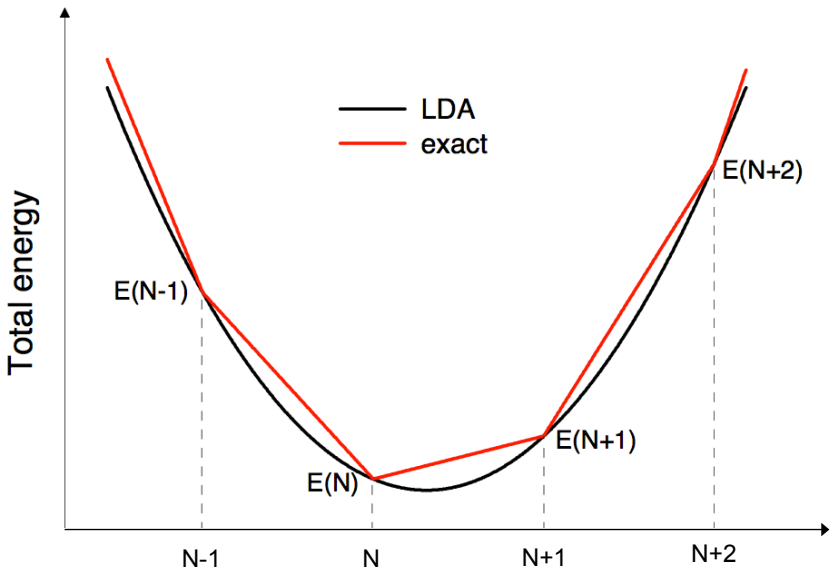


Fig. 2: *Sketch of the total energy profile as a function of number of electrons in a generic atomic system in contact with a reservoir of electrons. Adapted from [23].*

The Eqn. 61 is known as *Andersen's force theorem* and it's basic message is that in a linear approximation, the interaction drops out of the total energy change. The self-consistency step can therefore be ignored except for the surface term in the case of the volume change. Therefore, we came to a very important conclusion which can save a lot of computational effort: when there is a small perturbation in the system, during which there is no change of the volume, the difference in the total energies of the initial and the perturbed state can be substituted with the difference of the sums of eigenvalues of these two states.

4 Away from LDA and GGA

4.1 Self-interaction correction

It was noticed early on, that the Hartree electrostatic energy $E_{\rm es}$ given by Eq. 18 contains a contribution due to a spurious interaction of an electron with itself, i.e., the *self-interaction*. Naive thinking would lead us to a conclusion that self-interaction should be compensated by a part of the exchange-correlation energy in Eq. 31. This is indeed not so far from the truth: for example for a one-electron hydrogen atom which should have zero electron-electron interaction up to 95% of the electrostatic energy is compensated by exchange-correlation energy [24]. What do remaining 5% come from?

Imagine that our system is in contact with a reservoir of electrons. This means that our system can exchange an integer number of electrons with its environment at all times. A situation of a fractional number of electrons in our system can be described statistically, so that the total energy of a system with $N+\epsilon$ electrons, where $0<\epsilon<1$, can be described as [25]

$$E_{N+\epsilon} = (1-\epsilon)E_N + \epsilon E_{N+1}, \tag{62}$$

where E_N and E_{N+1} are the total energies of the system with integer number N and $N+1$

electrons. Similarly, the density of the system with $N + \epsilon$ electrons is given by:

$$n_{N+\epsilon} = (1 - \epsilon)\, n_N + \epsilon\, n_{N+1} \tag{63}$$

The total energy as a function of the occupation number constructed in such a way is a sequence of line segments of constant slope, Fig. 2, with the discontinuity of the slope and values of the slope from opposite sides of an integer occupation corresponding to the respective ionization potential. This is how the exact DFT total energy behaves.
Let us examine closer the case of a system with the fractional charge of $0 < \epsilon < 1$, which gives $n_\epsilon = \epsilon\, n_1$, where n_1 is the density of our system with one electron. For such a system there is no electron-electron interaction, which means that

$$E_{xc}[n_\epsilon] = -E_{\rm es}[n_\epsilon]. \tag{64}$$

Since $E_{\rm es}$ involves a product of the density with itself, clearly, the latter condition can be formulated as:

$$E_{xc}[\epsilon\, n_1] = -\epsilon^2\, E_{\rm es}[n_1]. \tag{65}$$

It can be shown that in LDA this exact equality is not satisfied [26]:

$$E_{xc}^{LDA}[\epsilon\, n_1] < -\epsilon^2\, E_{\rm es}[n_1]. \tag{66}$$

The approximative nature of LDA results in a total energy curve which, away from integer values of electrons, has an unphysical slope as a function of electron occupation (see for example Fig. 2, where a spurious minimum of the LDA curve is acquired far away from the corresponding minimum of the exact DFT total energy). The difference between the exact DFT curve and its LDA or GGA approximation stems from the *self-interaction error* and is at the core of many problems of the LDA, including the failure in description of strongly-correlated materials (see also further) [23].
Very often the difference between the exact DFT and the LDA total energy at fractional occupation is called the *delocalization error*, which underlines the tendency of the local density approximation to delocalize electrons in the system. To understand this better, consider an example of H_2^+ molecule. It is known that while the LDA describes the chemical bond rather well, it fails completely at a very large separation between the H atoms, resulting in a system with a very delocalized electron, which is effectively equivalent to two half-occupied H atoms at the dissociation limit, with a very low total energy. The reason for this is exactly the delocalization error, which pushes each single hydrogen atom toward the lower energy which is acquired at fractional occupation. We can thus say that the delocalization error tends to push the localized initially electrons away from the atoms on which they are localized, which results in lowering of the energy and electron delocalization.
Let us turn for a moment to the Hartree-Fock (HF) approximation. Within this approximation, we have orbitals $\psi_{i\sigma}$ with occupation numbers $f_{i\sigma}$ and spin $\sigma = (\uparrow, \downarrow)$ and the charge density is constructed as follows:

$$n(\mathbf{r}) = \sum_{i\sigma} n_{i\sigma}(\mathbf{r}) = \sum_{i\sigma} f_{i\sigma} |\psi_{i\sigma}(\mathbf{r})|^2. \tag{67}$$

In HF approximation, the energy of interaction between the electrons is given by the sum of two terms, the electrostatic energy Eq. 18, and an exchange energy:

$$E_x^{HF} = -\frac{1}{2} \sum_{ij\sigma} f_{i\sigma} f_{j\sigma} \int d\mathbf{r} \int d\mathbf{r}' \, \frac{\psi_{i\sigma}^*(\mathbf{r}) \psi_{j\sigma}^*(\mathbf{r}') \psi_{j\sigma}(\mathbf{r}) \psi_{i\sigma}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}. \tag{68}$$

Clearly, the $i = j$ terms in E_x^{HF} cancel completely the self-interaction part of $E_{\rm es}[n]$, which means that the HF approximation is self-interaction-free. And although the HF approximation works quite badly for solids, partially due to the fact that it neglects the correlation completely, its inspection leads us to a correct conclusion that in the DFT-based schemes presented above the only partial cancellation of self-interaction occurs due to approximations made on the exchange-correlation energy functional and utilization of the local one-electron potential in Kohn-Sham equations.

The most natural way to fight the self-interaction problems is to introduce an orbital-dependent exchange-correlation functional [15], which, given an approximation E_{xc}^0 to the true exchange-correlation functional (e.g. LDA), reads:

$$E_{xc}^{SIC} := E_{xc}^0[n^\uparrow, n^\downarrow] - \sum_{i\sigma} \delta_{i\sigma}, \tag{69}$$

where

$$\delta_{i\sigma} = E_{\rm es}[n_{i\sigma}] + E_{xc}^0[n_{i\sigma}, 0]. \tag{70}$$

The role of "parameters" $\delta_{i\sigma}$ is that of *self-interaction correction* of corresponding orbitals. Given an exact exchange-correlation functional the self-interaction correction (SIC) vanishes. The self-interaction correction from above can be also put into self-consistency with the corresponding Kohn-Sham equations in which the orbital-dependent effective potential reads:

$$V_{eff,i\sigma}^{SIC}(\mathbf{r}) = v\psi_s^\sigma(\mathbf{r}) - e^2 \int \frac{n_{i\sigma}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}' - \mu_{xc}^\uparrow[n_{i\sigma}, 0], \tag{71}$$

where v_s^σ is taken from Eq. 48. Note, that generally the SIC potential is not invariant with respect to the choice of the orbitals and thus the results of the SIC calculation can depend on this choice.

The SIC calculations have been intensively performed in the past and are performed very often also nowadays. The general conclusion is that although the values of the exchange and correlation energies, as well as corresponding exchange-correlation holes, see Fig. 1, are greatly improved upon considering SIC, this occurs mainly due to localized core electrons, which play a minor role for many processes taking place between valence electrons. This is due to the fact that the role of SIC becomes increasingly less important as the degree of localization of electrons decreases. One of the general observations is that the exchange-correlation potential becomes more attractive and the eigenenergies are lowered as the SIC is introduced. It can be also shown that the SIC eigenvalues are in much better agreement with experimental ionization energies, that the DFT eigenvalues computed without SIC [15]. Importantly, the band gaps in semiconductors and insulators are described substantially better with SIC.

4.2 LDA + U

One of the most "famous" failures of the LDA lies in description of strongly-correlated materials, which, besides delocalized s- and p-electrons, contain well-localized partially filled d- and f-shells. Very often treating both types of electrons present in such a material with a unified single-particle mean-field like LDA approach leads to rather bad results, as far as the single-particle energies, as well as spin and orbital moments are concerned. This may result even in an incorrect prediction of the metallic ground state for many insulators. Attempting to cure this problem via SIC-based methods described above often does not lead to essential improvements.

A commonly practiced way to fight this problem lies in adding an orbital-dependent single-particle potential to localized electrons subject to strong correlation, which includes explicitly the information about the strength of the Coulomb repulsion within the set of problematic states. Within the so-called LDA+U approach [27], we single out a subspace of states (say d-states for simplicity) which we want to treat in the Hubbard-like manner, i.e., we want to introduce a d-d Coulomb interaction via the Hamiltonian:

$$\frac{1}{2}U\sum_{i\neq j} n_i n_j, \tag{72}$$

where U is some parameter, the meaning of which we clarify later, and n_i is the occupation number of orbital i from our subspace. The total number of electrons in our subspace is $N = \sum_i n_i$. This type of Hamiltonian is able to account for many correlation effects which are missed within the LDA. This Hamiltonian gives the total Coulomb energy due to d-d interactions as

$$E_{dc} = \frac{U(N-1)N}{2}. \tag{73}$$

Since empirically we know that the total Coulomb energy is rather well predicted within the LDA, in order to introduce a total energy functional which depends on a set of $\{n_i\}$ in a Hubbard-like picture we have to subtract E_{dc} from the LDA total energy (thus the name double counting correction, "dc"), and construct the LDA+U total energy functional as follows:

$$E^{LDA+U}[n,\{n_i\}] = E^{LDA}[n] + \frac{1}{2}U\sum_{i\neq j} n_i n_j - \frac{U(N-1)N}{2}. \tag{74}$$

This total energy functional will give us correct total energy, but it will modify the single-particle energies, which are poorly reproduced with LDA. Indeed, the singe-particle energies can be readily found:

$$\varepsilon_i = \frac{\partial E^{LDA+U}}{\partial n_i} = \varepsilon_i^{LDA} + U\left(\frac{1}{2} - n_i\right), \tag{75}$$

reflecting a splitting proportional to U between the occupied and unoccupied states, in analogy to the lower and upper Hubbard bands within the Hubbard model. The corresponding single-particle effective potential is also modified to read:

$$V_i(\mathbf{r}) = V^{LDA}(\mathbf{r}) + U\left(\frac{1}{2} - n_i\right). \tag{76}$$

Let us try to give now a rigorous meaning to parameter U. Imagine that we are given a density matrix $\hat{n}^\sigma = \{n^\sigma_{mm'}\}$, corresponding to localized within atomic spheres d-states with m as the magnetic number within the d-subspace. Then, in analogy to previous considerations, the LDA+U total energy functional is written as:

$$E^{LDA+U}[n,\{n^\sigma_{mm'}\}] = E^{LDA}[n] + E^U[\{n^\sigma_{mm'}\}] - E_{dc}[\{n^\sigma_{mm'}\}]. \tag{77}$$

In the latter expression E^U is the rotationally invariant representation of the Hubbard-like Hamiltonian from above [28]:

$$E^U[\{n^\sigma_{mm'}\}] = \frac{1}{2}\sum_{all}\{\langle mm'' \,|\, V_{ee} \,|\, m'm'''\rangle\, n^\sigma_{mm'} n^{-\sigma}_{m''m'''} - \tag{78}$$

$$- \left[\langle mm'' \,|\, V_{ee} \,|\, m'm'''\rangle - \langle mm'' \,|\, V_{ee} \,|\, m'''m'\rangle\right] n^\sigma_{mm'} n^\sigma_{m''m'''}\}, \tag{79}$$

with the matrix elements of the screened Coulomb interaction V_{ee} among the d-electrons. The expression for the mean-field double counting corrections in the so-called fully localized limit reads:

$$E_{dc}[\{n^{\sigma}_{mm'}\}] = \frac{UN(N-1)}{2} - \frac{1}{2}J\left[N^{\uparrow}(N^{\uparrow}-1) + N^{\downarrow}(N^{\downarrow}-1)\right], \tag{80}$$

in which $N^{\sigma} = \mathrm{Tr}(n^{\sigma}_{mm'})$, $N = \sum_{\sigma} N^{\sigma}$, while U and J are the so-called screened Coulomb and exchange parameters. In order to determine the exact values of the matrix elements of V_{ee} we can assume a completely atomic-like point of view, i.e., since the screened Coulomb interaction potential can be written around atomic center as a linear combination of spherical harmonics, we can deduce that

$$\langle mm'' \,|\, V_{ee} \,|\, m'm''' \rangle = \sum_{l'=0}^{\infty} \frac{4\pi}{2l'+1} \sum_{p=-l'}^{+l'} \langle lm \,|\, Y_{l'p} \,|\, lm' \rangle \left\langle lm'' \,\middle|\, Y^{*}_{l'p} \,\middle|\, lm''' \right\rangle F^{l'}, \tag{81}$$

where F^{l} are the screened Slater integrals. For $3d$ electrons, for example, only F^0, F^2 and F^4 play a role, and the parameters U and J can be written as follows:

$$U = F^0, \quad J = \frac{F^2 + F^4}{14}. \tag{82}$$

The ratio of F^2/F^4 is almost constant among the $3d$ elements (≈ 0.625). The modifications in the total energy due to the LDA+U method lead also to the additional contribution to the single-particle effective potential, which, in addition to the LDA part, contains now also the following term:

$$\sum_{mm'} |lm\sigma\rangle \, V^{\sigma}_{mm'} \, \langle lm'\sigma| \,, \tag{83}$$

where $V^{\sigma}_{mm'}$ can be expressed in terms of $\{n^{\sigma}_{mm'}\}$, U, J, and matrix elements of V_{ee}. The LDA+U addition to the effective potential is manifestly a projection operator, which means that it depends not only on the density matrix but also on the exact shape of the orbitals chosen to be treated "differently" from itinerant mean-field like electrons. However, the dependence of the results on the choice of the orbitals is not as pronounced as one might expect. In practical applications of the described above scheme the parameters U and J can be computed from constrained calculations. Also, very often the parameters U and J are treated as independent parameters used to tune the electronic structure of a considered material to achieve an agreement with known *a priori* properties. In the latter case the matrix elements of V_{ee} are deduced from the density matrix and given U, J-parameters, which allows us to construct the single-particle potential together with E^U and E_{dc} contributions to the total energy.

4.3 Hybrid functionals

As we have seen previously, the Hartree-Fock approximation, although leading to non-local effective potential, is self-interaction free. In the course of the past years a technique for constructing functionals which combine a fraction of non-local HF exchange with local exchange-correlation functionals has been developed. Such "hybrid" functionals, besides been partially self-interaction free, were shown to be an essential improvement over LDA and GGA in describing the properties of band-gap materials. The deep justification for constructing the hybrid functionals lies in the adiabatic connection formula, which establishes the connection between

fully interacting case and the fully non-interacting Kohn-Sham system via tuning the interaction strength λ, as discussed around Eq. 39. In the limit of weak interaction the exchange energy becomes identical to the HF exchange energy given by Eq. 68, which suggests that a certain fraction of the HF exchange should be added to the LDA or GGA exchange-correlation functional, leading to the hybrid exchange-correlation functional:

$$E_{ex}^{HYB} = E_{xc}^{LDA} + \alpha \left(E_x^{HF} - E_x^{LDA} \right). \tag{84}$$

By assuming a certain perturbation-theory inspired shape for the adiabatic-connection integrand Perdew and co-workers deduced the optimal value for α of 0.25. In combination with the PBE local functional [29] the resulting hybrid functional constructed according to Eq. 84 is called PBE0 [30]. Further, Heyd and co-workers [31] attempted to decompose the unscreened Coulomb potential taking place in Eq. 68 into the short-ranged (SR) and long-ranged (LR) parts following representation:

$$v(r) = \frac{1}{r} = \frac{\mathrm{erf}(\omega r)}{r} + \frac{\mathrm{erfc}(\omega r)}{r} = v^{LR}(r) + v^{SR}(r), \tag{85}$$

where $\mathrm{erf}(x)$ is the error function, $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$, and ω is an adjustable screening parameter. The hybrid functional constructed following this decomposition is called HSE, and it reads:

$$E_{ex}^{HSE} = E_{xc}^{PBE} + \alpha \left(E_x^{HF,SR}(\omega) - E_x^{PBE,SR}(\omega) \right), \tag{86}$$

where the screened by an optimized value of $\omega = 0.15$ potential v^{SR} is used in the expression for the HF exchange in Eq. 68. $E_x^{PBE,SR}(\omega)$ is the local functional which can be obtained from the PBE procedure which takes v^{SR} as an input for interaction potential. The motivation behind the latter expression is due to an empirical observation that, given $E_x^{PBE} = E_x^{PBE,SR} + E_x^{PBE,LR}$, the long-ranged parts of E_x^{PBE} and E_x^{HF}, i.e. $E_x^{PBE,LR}$ and $E_x^{HF,LR}$ respectively, cancel each other [31]. Moreover, the application of the HSE hybrid functional proved to systematically improve the band gaps in semiconductors. One of the key features of the hybrid functionals is that, derived from generally non-local expression for the total energy functional as given above, the effective Kohn-Sham potential, in addition to the local part contains also a non-local contribution which presents a significant challenge to compute [32].

References

[1] L. Thomas, Proc. Cambridge Phil. Soc. **23**, 542 (1927).

[2] E. Fermi, Rend. Accad. Naz.Linzei **6**, 602 (1927).

[3] D. R. Hartree, Proc. Cambridge Philos. Soc. **24**, 89 (1928).

[4] J. Slater, *The Self-Sonsistent Field of Molecules and Solids* (McGraw-Hill, New York, 1974).

[5] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).

[6] W. Kohn and L. Sham, Phys. Rev. A **140**, 1133 (1965).

[7] O. Gunnarsson and H. Hjelmberg, Phys. Scr. **11**, 97 (1975).

[8] Gell-Mann and K. Brueckner, Phys. Rev. **106**, 369 (1957).

[9] E. Wiegner, Phys. Rev. **46**, 1002 (1934).

[10] D. Ceperly and B. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[11] R. Dreitzler and E. Gross, *Density Functional Theory* (Springer, Berlin, 1990).

[12] G. Srivastava, J. Phys. A **17**, L317 (1984).

[13] O. Eriksson, J. Trygg, O. Hjortstam, B. Johansson, and J. Wills, Surf. Sci. **382**, 93 (1997).

[14] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[15] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[16] J. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[17] N. Ashcroft and N. Mermin, *Solid State Physics* (Holt-Saunders Japan LTD., 1976).

[18] O. Jepsen and O. Andersen, Solid. State Commun. **9**, 1763 (1971).

[19] P. Blöchl, O. Jepsen, and O. Andersen, Phys. Rev. B **49**, 16223 (1994).

[20] M. Methfessel and T. Paxton, Phys. Rev. B **40**, 3616 (1989).

[21] A. Macintosh and O. Andersen, *Electrons at the Fermi surface* (edited by M. Springford, page 149, Cambridge University Press, 1980).

[22] M. Methfessel and J. Kübler, J. Phys. F: Metal Phys. **12**, 141 (1982).

[23] M. Cococcioni and S. D. Gironcoli, Phys. Rev. B **71**, 035105 (2005).

[24] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[25] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).

[26] Y. Zhang and W. Yang, J. Chem. Phys. **109**, 2604 (1998).

[27] V. I. Anisomov, F. Aryasetiawan, and A. I. Lichtenstein, J. Phys.: Condens. Matter **9**, 767 (1997).

[28] A. I. Lichtenstein, V. I. Anisomov, and J. Zaanen, Phys. Rev. B **52**, R5468 (1995).

[29] J. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[30] J. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys **105**, 9982 (1996).

[31] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys **118**, 8207 (2003).

[32] M. Schlipf, M. Betzinger, C. Friedrich, M. Ležaić, and S. Blügel, Phys. Rev. B **84**, 125142 (2011).

A 3 Hartree-Fock and Quantum Chemical Correlation Methods [1]

M. Betzinger

Peter Grünberg Institut

Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

The electronic properties of an atom, a molecule, or a solid are governed by the many electron Schrödinger equation

$$\hat{H}\Psi_k(\mathbf{x}_1, ..., \mathbf{x}_N) = E_k\Psi_k(\mathbf{x}_1, \ldots, \mathbf{x}_N) \tag{1}$$

with the N-electron wave function Ψ_k of quantum number k and energy E_k. Each of the N wave-function arguments $\mathbf{x}_i$ $(i = 1, \ldots, N)$ subsumes space and spin coordinate $\mathbf{x}_i = (\mathbf{r}_i, s_i)$. For the Hamiltonian $\hat{H}$ we employ the Born-Oppenheimer approximation [1], i.e.,

$$\hat{H} = \sum_{i=1}^{N}\left(-\frac{1}{2}\nabla_i^2 - \sum_{\alpha}\frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|}\right) + \frac{1}{2}\sum_{\substack{i,j \\ i\neq j}}^{N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}\,. \tag{2}$$

The positions of the atomic nuclei $\mathbf{R}_\alpha$ enter the electronic Hamiltonian $\hat{H}$ solely as parameters. Their positive point-charges Z_α give rise to an attractive external potential, in which the electrons move. In addition, the electrons interact with each other via the electrostatic Coulomb interaction. The mutual Coulomb repulsion of the electrons leads to a correlated electron motion, i.e., the motion of an individual electron is influenced by all other electrons and conversely it affects the motion of all others electrons. In fact, it is the Coulomb interaction which hampers and even impedes a direct solution of the Schrödinger equation apart from a few special cases.
In density functional theory a direct solution of the many-electron Schrödinger equation is avoided by mapping the interacting electron system onto an artificial non-interacting system whose ground-state electron density coincides by construction with that of the true fully-interacting electron system (for details see lecture A2 of M. Ležaić).
Wave-function based quantum chemical methods pursue a different approach: instead of mapping the interacting Hamiltonian onto a much simpler and solvable auxiliary system they use the exact many electron Hamiltonian (in the Born-Oppenheimer approximation) and employ approximations for the many electron wave function $\Psi_k(\mathbf{x}_1, \ldots, \mathbf{x}_N)$. The simplest of these quantum chemical methods is the Hartree-Fock approach [2], which approximates the many electron wave function as a single Slater determinant, i.e., a sum of products of one-particle orbitals obeying the Pauli principle. The one-particle orbitals are chosen such that the total energy, the expectation value of the Hamiltonian with the Slater determinant, is minimized giving rise to a Schrödinger-like equation for the orbitals. Since the Hartree-Fock method is a variational approach, its energy is a strict upper bound for the true ground-state total energy of the system. The Hartree-Fock approach forms the basis for a series of more advanced and more accurate quantum chemical methods, for example Møller-Plesset (MP) perturbation theory [3], the configuration interaction (CI) approach [4], or the coupled cluster (CC) method [5]. MP perturbation theory is based on the Hartree-Fock Hamiltonian and the corresponding Slater determinant. The difference between the true many electron Hamiltonian $\hat{H}$ and the Fock Hamiltonian is treated as a perturbation. The Hartree-Fock energy and orbitals are improved order-by-order employing Rayleigh-Schrödinger perturbation theory. In contrast, CI and CC are not perturbative approaches. In the CI method, the many-electron wave function is expanded as a linear combination of Slater determinants, that are usually formed from the set of occupied and unoccupied Hartree-Fock orbitals. By taking into account in principle all possible Slater determinants in this expansion CI would yield the exact solution of the problem at hand. In practice, of course, only a finite number of Slater determinants can be considered leading to an approximate, but systematically improvable, solution of the many electron problem. Similar to the CI

approach, the CC method writes the many-electron wave function as a superposition of Slater determinants, but uses an exponential instead of a linear ansatz (as will be explained later).
The aim of this chapter is to give an introduction to the Hartree-Fock method, MP perturbation theory, the CI and CC approach. The main ideas and basic concepts of the different approaches should be conveyed. The interested reader is referred to the literature for an in depth discussion of these approaches. The book of Jensen [6], Szabo and Ostlund [7] as well as the review articles [8, 9, 10, 11] have served as a basis for this chapter.

2 The electronic wave function

Besides that the N electron wave function $\Psi(\mathbf{x}_1, ..., \mathbf{x}_N)$ solves the Schrödinger equation [Eq. (1)] it has to fulfill additional requirements.[2] First of all, the wave function must be square integrable. Hence, it can be normalized and $|\Psi(\mathbf{x}_1, ..., \mathbf{x}_N)|^2$ corresponds to the probability distribution of finding the N electrons at the positions $\mathbf{r}_1, ..., \mathbf{r}_N$ with spin coordinate $s_1, \ldots, s_N$. From the indistinguishability of the electrons follows that $|\Psi(\mathbf{x}_1, ..., \mathbf{x}_N)|^2$ is the probability of finding any of the N electrons at position $\mathbf{r}_1$ with spin s_1, any of the remaining $N-1$ electrons at $\mathbf{r}_2$ with s_2 and so on and so forth. The indistinguishability of the electrons combined with their fermionic character leads to the Pauli principle. It requires that the wave function $\Psi(\mathbf{x}_1, ..., \mathbf{x}_N)$ is antisymmetric with respect to permutations of (any) two electrons

$$\Psi(\mathbf{x}_1, ..., \mathbf{x}_i, ..., \mathbf{x}_j, ..., \mathbf{x}_N) = -\Psi(\mathbf{x}_1, ..., \mathbf{x}_j, ..., \mathbf{x}_i, ..., \mathbf{x}_N)\,. \tag{3}$$

Since the Hamilton operator $\hat{H}$ is spin independent, it commutes with the z-component $\hat{S}_z = \sum_i^N \hat{S}_z^i$ and the square $\hat{S}^2 = (\sum_i^N \hat{S}^i)^2$ of the total electron spin operator, where $\hat{S}^i$ denotes the spin operator of the i-th electron. Consequently, the proper N electron wave function has to be square integrable, antisymmetric with respect to permutations of two electrons, and an eigenfunction of $\hat{S}_z$ and $\hat{S}^2$. When constructing an approximate wave function these conditions should be taken into account as constraints.

2.1 Representation by Slater determinants

If we disregard the Coulomb interaction between the electrons for a moment the many electron Hamiltonian becomes a sum of one-particle Hamiltonians $\hat{h}(\mathbf{r})$

$$\hat{H}_0 = \sum_i \hat{h}(\mathbf{r}_i) \tag{4}$$

with $\hat{h}(\mathbf{r}_i) = -\frac{1}{2}\nabla_i^2 - \sum_\alpha \frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|}$. (The subscript 0 indicates that the electron-electron interaction has been neglected.) In this simplified case, a product $\Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N) = \varphi_i(\mathbf{x}_1)\varphi_j(\mathbf{x}_2) \ldots \varphi_N(\mathbf{x}_N)$ of one-particle spin functions $\varphi_i(\mathbf{x}) = \varphi_i(\mathbf{r})\chi_i(s)$, each obeying

$$\hat{h}(\mathbf{r})\varphi_i(\mathbf{x}) = \epsilon_i\varphi_i(\mathbf{x}) \tag{5}$$

$$\hat{S}_z\varphi_i(\mathbf{x}) = \sigma_i\varphi_i(\mathbf{x}) \tag{6}$$

$$\hat{S}^2\varphi_i(\mathbf{x}) = \frac{3}{4}\varphi_i(\mathbf{x})\ , \tag{7}$$

[2] The quantum number k of the N electron wave function is suppressed in the following.

is an eigenfunction of the Hamiltonian $\hat{H}_0$. ($\chi_i(s)$ denotes a two component spin eigenfunction of $\hat{S}^2$ and $\hat{S}_z$ with eigenvalue $3/4$ and $\sigma_i = \pm 1/2$, respectively.) Under the premise that the one-particle functions are square integrable, the product function is square integrable as well. However, the simple product ansatz violates the Pauli principle. A properly antisymmetric wave function Ψ_0 can be constructed from the one-particle functions through the determinant

$$\Psi_0(\mathbf{x}_1, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_i(\mathbf{x}_1) & \varphi_i(\mathbf{x}_2) & \ldots & \varphi_i(\mathbf{x}_N) \\ \varphi_j(\mathbf{x}_1) & \varphi_j(\mathbf{x}_2) & \ldots & \varphi_j(\mathbf{x}_N) \\ \vdots & \vdots & & \vdots \\ \varphi_N(\mathbf{x}_1) & \varphi_N(\mathbf{x}_2) & \ldots & \varphi_N(\mathbf{x}_N) \end{vmatrix}, \tag{8}$$

where the factor $1/\sqrt{N!}$ guarantees the normalization of the wave function. Such a wave function is called a Slater determinant. It changes its sign if two electrons are permutated. Moreover, the determinant vanishes if two electrons occupy the same orbital. As a shorthand notation for the Slater determinant we introduce the abstract notation $|\varphi_i \varphi_j \ldots \varphi_N\rangle$. If we refer to a simple product of one-particle orbitals without symmetrization we write $|ij \ldots N\rangle$.
While Slater determinants formed from one-particle spin orbitals are eigenfunctions of $\hat{S}_z$ with eigenvalue $M = \frac{1}{2}(N_\uparrow - N_\downarrow)$, where $N_\uparrow$ and $N_\downarrow$ denote the number of spin-up and -down electrons, respectively, they are in general no eigenfunctions of the operator $\hat{S}^2$. However, one can build a linear combination of Slater determinants with the same spatial orbitals but different spin configurations such that the sum is an eigenfunction of $\hat{S}^2$ [12]. These linear combinations are called spin-adapted Slater determinants or configuration state functions. For the rest of the manuscript we restrict ourselves to the simple case of an even number of electrons N, where the $N/2 = N_\uparrow = N_\downarrow$ spatial orbitals are doubly occupied. In this special case, the single Slater determinant is a eigenfunction of $\hat{S}^2$ with eigenvalue $S = 0$, i.e., it corresponds to a singlet state.
At first glance the construction of a (single) Slater determinant from one-particle spin functions seems of very little practical use since it is only a solution of the many electron Hamiltonian if the electron-electron interaction is neglected. However, provided that the set of one-particle functions $\{\varphi_p(\mathbf{x})\}$ is complete, the space spanned by all possible Slater determinants constructed from these one-particle orbitals forms a complete basis for expanding the antisymmetric, interacting N electron wave function $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)$. In other words, each interacting many-electron wave function can be represented exactly by a sum over Slater determinants, if the one-particle orbitals used to construct the Slater determinants are complete.

2.2 Second Quantization

The calculation of Hamiltonian matrix elements between Slater determinants using its basic definition [Eq. (8)] is a laborious task, but in principle possible. Yet, it can be significantly facilitated by exploiting the formalism of second quantization. The latter has been already introduced and discussed in the lecture B1 of E. Koch. Here, we briefly summarize the main aspects necessary for deriving and understanding wave-function based quantum chemical approaches.
The many-electron Hamiltonian $\hat{H}$ in second quantization becomes

$$\hat{H} = \sum_{pq} h_{pq} \hat{c}_p^\dagger \hat{c}_q + \frac{1}{2} \sum_{pqrs} \langle pq|rs\rangle \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_s \hat{c}_r \,, \tag{9}$$

where the sums run over a complete set of one-electron spin orbitals $\{\varphi_p(\mathbf{x})\}$, h_{pq} denotes the matrix element $\langle p|\hat{h}|q\rangle = \langle\varphi_p|\hat{h}|\varphi_q\rangle = \int \varphi_p^*(\mathbf{x}_1)h(\mathbf{r}_1)\varphi_q(\mathbf{x}_1)d\mathbf{x}_1$, $\langle pq|rs\rangle$ is the two-electron integral defined by $\langle pq|rs\rangle = \iint \frac{\varphi_p^*(\mathbf{x}_1)\varphi_q^*(\mathbf{x}_2)\varphi_r(\mathbf{x}_1)\varphi_s(\mathbf{x}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d\mathbf{x}_1 d\mathbf{x}_2$, and $\hat{c}_p^\dagger$ and $\hat{c}_p$ are so-called creation and annihilation operators, respectively. For simplicity, we require that the one-electron spin orbitals are orthogonal.
We define a creation operator by its action on a Slater determinant

$$\hat{c}_p^\dagger|\varphi_i\varphi_j\ldots\varphi_N\rangle = |\varphi_p\varphi_i\varphi_j\ldots\varphi_N\rangle. \tag{10}$$

While in the initial determinant $|\varphi_i\varphi_j\ldots\varphi_N\rangle$ the N electrons occupy the one-particle spin orbitals φ_i to φ_N, $\hat{c}_p^\dagger|\varphi_i\varphi_j\ldots\varphi_N\rangle$ describes an $N+1$ electron Slater determinant, where additionally the orbital φ_p is occupied.[3] In a similar manner, the annihilation operator $\hat{c}_p$ removes an electron from the orbital φ_p

$$\hat{c}_p|\varphi_p\varphi_i\varphi_j\ldots\varphi_N\rangle = |\varphi_i\varphi_j\ldots\varphi_N\rangle. \tag{11}$$

By defining the vacuum as a state $|0\rangle$ containing no electrons, the Slater determinant $|\varphi_i\varphi_j\ldots\varphi_N\rangle$ becomes

$$|\varphi_i\varphi_j\ldots\varphi_N\rangle = \hat{c}_i^\dagger\hat{c}_j^\dagger\ldots\hat{c}_N^\dagger|0\rangle\,. \tag{12}$$

Acting with the annihilation operator on the vacuum gives rise to

$$\hat{c}_p|0\rangle = 0\,. \tag{13}$$

The creation and annihilation operators obey the anticommutation relations

$$\hat{c}_p^\dagger\hat{c}_q^\dagger + \hat{c}_q^\dagger\hat{c}_p^\dagger = 0 \tag{14}$$
$$\hat{c}_p\hat{c}_q + \hat{c}_q\hat{c}_p = 0 \tag{15}$$
$$\hat{c}_p^\dagger\hat{c}_q + \hat{c}_q\hat{c}_p^\dagger = \delta_{pq} \tag{16}$$

where δ_{pq} is the Kronecker delta function, which equals 1 if $p = q$ and is 0 for $p \neq q$. These commutation rules ensure the correct antisymmetry of the wave function, for example

$$|\varphi_j\varphi_i\ldots\varphi_N\rangle = \hat{c}_j^\dagger\hat{c}_i^\dagger\ldots\hat{c}_N^\dagger|0\rangle = -\hat{c}_i^\dagger\hat{c}_j^\dagger\ldots\hat{c}_N^\dagger|0\rangle = -|\varphi_i\varphi_j\ldots\varphi_N\rangle\,. \tag{17}$$

Furthermore, these rules assure that two electrons do not occupy the same spin orbital twice as from the anticommutation rule [Eq. (14)] follows for $p = q$

$$\hat{c}_p^\dagger\hat{c}_p^\dagger = -\hat{c}_p^\dagger\hat{c}_p^\dagger = 0\,. \tag{18}$$

In order to demonstrate the use of the second quantization formalism, we explicitly show how to calculate the matrix element of the Hamiltonian with the Slater determinant $|\Psi_0\rangle = |\varphi_i\varphi_j\ldots\varphi_N\rangle$. For the matrix element arising from the one-particle part of the second-quantized Hamiltonian we obtain

$$\langle\Psi_0|\sum_{pq} h_{pq}\hat{c}_p^\dagger\hat{c}_q|\Psi_0\rangle = \sum_{pq} h_{pq}\langle\varphi_i\varphi_j\ldots\varphi_N|\hat{c}_p^\dagger\hat{c}_q|\varphi_i\varphi_j\ldots\varphi_N\rangle\,. \tag{19}$$

The operator $\hat{c}_q$ tries to depopulate the orbital φ_q. If the orbital is not occupied in the initial wave function Ψ_0, it gives zero. The creation operator $\hat{c}_p^\dagger$ can only populate an orbital which is

[3] It is implicitly assumed that the orbital φ_p is initially unoccupied.

unoccupied before. Hence, $\hat{c}_p^\dagger$ acting on $\hat{c}_q|\varphi_i\varphi_j\ldots\varphi_N\rangle$ is only nonzero if $p \neq i, j, ..., q-1, q+1, ..., N$. Furthermore, the overlap of two Slater determinants is zero, if the Slater determinants differ in at least one orbital. A property resulting from the orthogonality of the one-particle spin orbitals. Combining these arguments finally yields

$$\langle\Psi_0|\sum_{pq} h_{pq}\hat{c}_p^\dagger\hat{c}_q|\Psi_0\rangle = \sum_i^{\text{occ.}} h_{ii}\,. \tag{20}$$

The two-electron part of the matrix element $\langle\Psi_0|\hat{H}|\Psi_0\rangle$ involves the elements

$$\langle\varphi_i\varphi_j\ldots\varphi_N|\hat{c}_p^\dagger\hat{c}_q^\dagger\hat{c}_s\hat{c}_r|\varphi_i\varphi_j\ldots\varphi_N\rangle\,. \tag{21}$$

Acting with the creation operators to the left and using the orthogonality of the Slater determinants we obtain

$$\begin{aligned}\langle\varphi_i\varphi_j\ldots\varphi_N|\hat{c}_p^\dagger\hat{c}_q^\dagger\hat{c}_s\hat{c}_r|\varphi_i\varphi_j\ldots\varphi_N\rangle &= \langle\hat{c}_q\hat{c}_p\varphi_i\varphi_j\ldots\varphi_N|\hat{c}_s\hat{c}_r\varphi_i\varphi_j\ldots\varphi_N\rangle &(22)\\ &= n_s n_r(1-\delta_{rs})(\delta_{rp}\delta_{sq}-\delta_{rq}\delta_{sp})\,, &(23)\end{aligned}$$

where n_i is the occupation number of the orbital i. It is 1 if the orbital is occupied and zero otherwise. The expectation value of the Hamiltonian $\hat{H}$ with the Slater determinant $|\Psi_0\rangle$ is then given by

$$\langle\Psi_0|\hat{H}|\Psi_0\rangle = \sum_i^{\text{occ.}} h_{ii} + \frac{1}{2}\sum_i^{\text{occ.}}\sum_{j\neq i}^{\text{occ.}}(\langle ij|ij\rangle - \langle ij|ji\rangle)\,. \tag{24}$$

3 Hartree-Fock method

In the Hartree-Fock method [2] the many electron ground-state wave function $\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)$ is approximated by the *best* single Slater determinant Ψ^{HF}. In this context *best* means that the orbitals forming the Hartree-Fock single Slater determinant are determined such that the total energy $E^{\text{HF}} = \langle\Psi^{\text{HF}}|H|\Psi^{\text{HF}}\rangle$ is minimized. Using the results of the previous section, the Hartree-Fock energy becomes a functional of the spin-orbitals $\{\varphi_i\}$

$$E^{\text{HF}}[\{\varphi_i\}] = \sum_{i=1}^{\text{occ.}} h_{ii} + \frac{1}{2}\sum_i^{\text{occ.}}\sum_{j\neq i}^{\text{occ.}}(\langle ij|ij\rangle - \langle ij|ji\rangle). \tag{25}$$

We search for those spin-orbitals $\{\varphi_i\}$ which minimize the Hartree-Fock energy subjected to the constraint that the orbitals remain orthonormal. In fact, this condition results in two requirements for each combination of orbitals i and j

$$\begin{aligned}\text{Re}(\langle\varphi_i|\varphi_j\rangle) - \delta_{ij} &= 0 &(26)\\ \text{Im}(\langle\varphi_i|\varphi_j\rangle) &= 0\,, &(27)\end{aligned}$$

which are taken into account by the technique of Lagrange's multipliers giving rise to the modified functional

$$\mathcal{L}[\{\varphi_i\},\{a_{ij}\},\{b_{ij}\}] = E^{\text{HF}}[\{\varphi_i\}] + \sum_i^{\text{occ.}}\sum_{j\leq i}\{a_{ij}[\text{Re}(\langle\varphi_i|\varphi_j\rangle)-\delta_{ij}] + b_{ij}\text{Im}(\langle\varphi_i|\varphi_j\rangle)\}\,. \tag{28}$$

The set of (real) Lagrange multipliers $\{a_{ij}\}$ and $\{b_{ij}\}$ guarantee the orthogonality and normalization of the orbitals. To account for each constraint only once, the second summation in Eq. (28) is restricted to $j \leq i$. Using the explicit expressions for the real and imaginary part in terms of $\langle\varphi_i|\varphi_j\rangle$ and $\langle\varphi_j|\varphi_i\rangle$ we end up with

$$\mathcal{L}[\{\varphi_i\},\{a_{ij}\},\{b_{ij}\}] \tag{29}$$

$$= E^{\mathrm{HF}}[\{\varphi_i\}] + \sum_i^{\mathrm{occ.}} \sum_{j\leq i} \frac{a_{ij} - ib_{ij}}{2}[\langle\varphi_i|\varphi_j\rangle - \delta_{ij}] + \frac{a_{ij} + b_{ij}}{2}[\langle\varphi_j|\varphi_i\rangle - \delta_{ij}] \tag{30}$$

$$= E^{\mathrm{HF}}[\{\varphi_i\}] + \sum_{i,j}^{\mathrm{occ.}} \epsilon_{ij}[\langle\varphi_i|\varphi_j\rangle - \delta_{ij}]\,, \tag{31}$$

where we defined the hermitian matrix of complex Lagrange multipliers $\epsilon_{ij} = \frac{1}{2}(a_{ij} - ib_{ij})$. We minimize $\mathcal{L}$ with respect to the wave function $\varphi_a^*(\mathbf{x})$ by setting the functional derivative $\delta\mathcal{L}/\delta\varphi_a^*(\mathbf{x})$ to zero

$$\frac{\delta\mathcal{L}}{\delta\varphi_a^*(\mathbf{x})} = \hat{h}(\mathbf{r})\varphi_a(\mathbf{x}) + \sum_{j\neq a}^{\mathrm{occ.}} \int \frac{\varphi_j^*(\mathbf{x}')\varphi_j(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}'\varphi_a(\mathbf{x}) \tag{32}$$

$$- \sum_{j\neq a}^{\mathrm{occ.}} \int \frac{\varphi_j^*(\mathbf{x}')\varphi_a(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}'\varphi_j(\mathbf{x}) - \sum_j \epsilon_{aj}\varphi_j(\mathbf{x}) = 0\,. \tag{33}$$

We note that differentiation with respect to $\varphi_a(\mathbf{x})$ instead of $\varphi_a^*(\mathbf{x})$ leads to the complex conjugate equation. The restriction in the summation $j \neq a$ in the first and second sum can be abandoned since the terms for $j = a$ cancel mutually. Then, the second term

$$\sum_j^{\mathrm{occ.}} \int \frac{\varphi_j^*(\mathbf{x}')\varphi_j(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}'\varphi_a(\mathbf{x}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}'\varphi_a(\mathbf{x}) = \hat{V}_{\mathrm{H}}(\mathbf{r})\varphi_a(\mathbf{x}) \tag{34}$$

corresponds to the classical Coulomb or Hartree term, i.e., the electrostatic potential generated by the charge $n(\mathbf{r}) = \sum_j^{\mathrm{occ.}} \sum_s |\varphi_j(\mathbf{r},s)|^2$ of all electrons of the system multiplied with the orbital $\varphi_a(\mathbf{x})$. In contrast, the third term has no classical interpretation. The latter arises solely from the antisymmetry of the wave function and is known as the non-local Hartree-Fock exchange potential

$$\hat{V}_{\mathrm{HF}}(\mathbf{x})\varphi_a(\mathbf{x}) = -\sum_j^{\mathrm{occ.}} \int \frac{\varphi_j^*(\mathbf{x}')\varphi_a(\mathbf{x}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{x}'\varphi_j(\mathbf{x})\,. \tag{35}$$

It is a non-local potential as it depends on the value of the unknown wave function $\varphi_a(\mathbf{x}')$ throughout all space. With these definitions Eq. (32) turns into

$$[\hat{h}(\mathbf{r}) + \hat{V}_{\mathrm{H}}(\mathbf{r}) + \hat{V}_{\mathrm{HF}}(\mathbf{x})]\varphi_a(\mathbf{x}) = \sum_j \epsilon_{aj}\varphi_j(\mathbf{x})\,. \tag{36}$$

This equation deviates from the form of a conventional Schrödinger equation through the sum on the right hand side. However, the Hartree-Fock total energy is invariant with respect to a unitary transformation of the orbitals: having found one set of orbitals $\{\varphi_i(\mathbf{x});\, i = 1,\ldots,N\}$

that solves Eq. (36) all set of functions $\{\varphi_i'(\mathbf{x}); i = 1, \ldots, N\}$ emerging from the initial set of orbitals by a unitary transformation U

$$\varphi_i'(\mathbf{x}) = \sum_a U_{ia}\varphi_a(\mathbf{x}) \tag{37}$$

are again solutions. In particular, we can choose U to be the unitary matrix that diagonalizes the matrix ϵ_{ij} of Lagrange multipliers. Since the latter is hermitian, this specific unitary matrix must always exist. Consequently, a unique set of spin orbitals $\varphi_j'(\mathbf{x})$ exists for which the matrix of Lagrange multipliers is diagonal and which fulfills

$$[\hat{h}(\mathbf{r}) + \hat{V}_{\mathrm{H}}(\mathbf{r}) + \hat{V}_{\mathrm{HF}}(\mathbf{x})]\varphi_j'(\mathbf{x}) = \epsilon_j'\varphi_j'(\mathbf{x})\,. \tag{38}$$

This set of spin orbitals is usually called canonical spin orbitals.
In the following, we omit the primes and simply write the Hartree-Fock equations as

$$\hat{f}(\mathbf{x})\varphi_j(\mathbf{x}) = \epsilon_j\varphi_j(\mathbf{x})\,, \tag{39}$$

where we additionally introduced the Fock operator $\hat{f}(\mathbf{x}) = \hat{h}(\mathbf{r}) + \hat{V}_{\mathrm{H}}(\mathbf{r}) + \hat{V}_{\mathrm{HF}}(\mathbf{x})$. The non-local Hartree-Fock exchange potential $\hat{V}_{\mathrm{HF}}(\mathbf{x})$, as indicated by the argument $\mathbf{x}$, depends in general on the spin coordinate s. By separating the spin-quantum number σ, which corresponds to spin-up ($\sigma = 1/2$) or spin-down ($\sigma = -1/2$), from the complete tuple of quantum numbers j, we can integrate out the spin-coordinates in Eq. (39) and obtain separate canonical Hartree-Fock equations for spin-up and -down

$$\hat{f}^\sigma(\mathbf{r})\varphi_j^\sigma(\mathbf{r}) = \epsilon_j^\sigma\varphi_j^\sigma(\mathbf{r}) \tag{40}$$

with $\hat{f}^\sigma(\mathbf{r}) = \hat{h}(\mathbf{r}) + \hat{V}_{\mathrm{H}}(\mathbf{r}) + \hat{V}_{\mathrm{HF}}^\sigma(\mathbf{r})$ and $\hat{V}_{\mathrm{HF}}^\sigma(\mathbf{r})\varphi_j^\sigma(\mathbf{r}) = -\sum_i^{\mathrm{occ.}} \int \frac{\varphi_i^{\sigma *}(\mathbf{r}')\varphi_j^\sigma(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}'\varphi_i^\sigma(\mathbf{r})$. As we restricted ourselves to the case of an even number of electrons N, where each spatial orbital is occupied twice by a spin-up and -down electron, we omit the spin index σ in the following.
Due to the similarity of the Hartree-Fock equation (39) with a Schrödinger equation we can identify the Lagrange multiplier ϵ_j as the orbital energy of the effective one-particle Hamiltonian. Please note that the sum of all occupied one-particle energies ϵ_j does not equal the Hartree-Fock total energy E^{HF}. For the orbital energy ϵ_j we find

$$\epsilon_j = \langle\varphi_j|\hat{h}|\varphi_j\rangle + \sum_{i=1}^{\mathrm{occ.}}\langle ji|ji\rangle - \sum_{i=1}^{\mathrm{occ.}}\langle ji|ij\rangle \tag{41}$$

and by comparing with Eq. (25) it is evident that

$$\sum_{j=1}^{\mathrm{occ.}} \epsilon_j \neq E^{\mathrm{HF}}\,. \tag{42}$$

In particular, the sum over the orbital energies counts the Coulomb and exchange energy twice. In addition to the purely formal identification of the Lagrange multipliers ϵ_j as orbital energies strict physical meaning can be attributed to them by Koopmans' theorem [13]. In order to derive Koopmans' theorem we deal with the question what is the ionization energy for removing an electron from the occupied orbital a in the Hartree-Fock method. This ionization energy

corresponds to the energy difference between the ground-state Hartree-Fock total energy of the $N-1$ and N electron system

$$\mathrm{I}_a = E_a^{\mathrm{HF}}(N-1) - E^{\mathrm{HF}}(N)\,, \tag{43}$$

where the subscript a indicates that an electron from the orbital a has been removed. In principle, the calculation of the ionization energy I_a requires two independent calculations for the N and $N-1$ electron system. Under the premise that the spin orbitals of the $N-1$ electron system are identical to those of the N electron system, the energy of the $N-1$ electron system is directly accessible

$$E_a^{\mathrm{HF}}(N-1) = \sum_{\substack{i=1\\ i\neq a}}^{\mathrm{occ.}} \langle\varphi_i|\hat{h}|\varphi_i\rangle + \frac{1}{2}\sum_{\substack{i=1\\ i\neq a}}^{\mathrm{occ.}}\sum_{j\neq i}^{\mathrm{occ.}} (\langle ij|ij\rangle - \langle ij|ji\rangle)\,, \tag{44}$$

where $\{\varphi_i\}$ are the orbitals from the N electron Hartree-Fock calculation. Forming the difference between Eq. (44) and (25) gives rise to the ionization energy

$$\mathrm{I}_a = -\epsilon_a\,. \tag{45}$$

Thus, the negative orbital energy corresponds to the energy required for removing an electron from that spin orbital. Analogously, the energy for adding an electron to the unoccupied spin orbital r, i.e., the electron affinity

$$\mathrm{A}_r = E^{\mathrm{HF}}(N) - E_r^{\mathrm{HF}}(N+1)\,, \tag{46}$$

is given by the negative energy of the unoccupied orbital

$$\mathrm{A}_r = -\epsilon_r. \tag{47}$$

In conclusion, under the assumption that the spin orbitals do not relax, if an electron is removed from or added to the N electron system, the negative of the occupied and unoccupied Hartree-Fock orbital energies can be identified as ionization energies and electron affinities, respectively.

A direct numerical solution of the (canonical) Hartree-Fock integro-differential equation [Eq. (39)] is possible for atoms. For molecules, clusters, and solids one usually introduces a basis set $\{\phi_\mu;\ \mu = 1,\ldots,N_{\mathrm{basis}}\}$ and expands the orbitals in this basis

$$\varphi_j(\mathbf{r}) = \sum_\mu C_{\mu j}\phi_\mu(\mathbf{r})\,. \tag{48}$$

In this way, the integro-differential equation turns into an algebraic equation for the expansion coefficients $C_{\mu j}$

$$\sum_\nu F_{\mu\nu}C_{\nu j} = \epsilon_j \sum S_{\mu\nu}C_{\nu j} \tag{49}$$

with the Fock matrix $F_{\mu\nu} = \langle\phi_\mu|\hat{f}|\phi_\nu\rangle$ and the overlap matrix $S_{\mu\nu} = \langle\phi_\mu|\phi_\nu\rangle$ of dimension $N_{\mathrm{basis}} \times N_{\mathrm{basis}}$. Since the Fock matrix elements $F_{\mu\nu}$ already depend on the orbitals (and thus on the coefficient matrix $C_{\mu j}$), equation (49) poses a non-linear problem, which is usually solved iteratively. Starting with an initial guess for the orbitals $\{\varphi_j\}$ the Fock matrix elements and the overlap matrix are calculated. Then Eq. (49), which corresponds to a generalized eigenvalue

problem, is solved by conventional numerical algebra tools giving rise to (intermediate) Hartree-Fock eigenvalues ϵ_j and the coefficient matrix $C_{\mu j}$. The latter corresponds to a new set of orbitals $\{\varphi_j\}$, with which the whole procedure is iterated until convergence in the orbitals is achieved. Usually, the density matrix

$$P_{\mu\nu} = \sum_{j=1}^{\text{occ.}} C_{\mu j} C_{\nu j}^{*} \tag{50}$$

is used as a measure of convergence. If self-consistency between input and output density matrix is achieved (within a specified criterion), we found a solution of the Hartree-Fock equations in the given basis set. In fact, we obtain N_{basis} Hartree-Fock orbitals and eigenvalues, from which the N energetically lowest states are occupied. If the basis set approaches completeness the orbitals and eigenvalues will approach the exact solution.
The computationally most demanding step in a practical calculation is typically the computation of the matrix elements of the non-local exchange potential

$$\langle \phi_\mu | \hat{V}_{\text{HF}} | \phi_\nu \rangle = -\sum_{\zeta\tau} P_{\tau\zeta} \langle \phi_\mu \phi_\zeta | \phi_\tau \phi_\nu \rangle \tag{51}$$

which formally scales as N_{basis}^4. Over the years several techniques, e.g., integral pre-screening, resolution of the identity [14, 15], have been developed to speed up the calculation.
Until now we have not specified the form of the basis functions $\phi_\mu(\mathbf{r})$. It can be Slater-type or Gaussian orbitals, which are normally employed in quantum chemical calculations for finite systems (atoms, molecules, etc.). For the calculation of solids, i.e., a periodic infinite crystal, plane waves are more suitable, for example.

4 Post-Hartree-Fock: Correlated methods

Electron correlation describes the phenomenon that the motion of the electrons constituting an atom, molecule, or solid is not independent. Two effects, the Pauli principle and the Coulomb interaction between the electrons, give rise to a correlated electron motion. While the former is often referred to as Fermi correlation, the latter is usually called Coulomb correlation.
As a measure of correlation we define the difference between the two-particle density $n(\mathbf{x}, \mathbf{x}')$, which describes the probability of finding an electron at $\mathbf{r}'$ with spin coordinate s' and another one at position $\mathbf{r}$ with spin coordinate s [$\mathbf{x} = (\mathbf{r}, s)$], and the product $n(\mathbf{x})n(\mathbf{x}')$ of one-particle densities

$$\Delta n(\mathbf{x}, \mathbf{x}') = n(\mathbf{x}, \mathbf{x}') - n(\mathbf{x})n(\mathbf{x}') \,. \tag{52}$$

For independent electrons the two-particle density $n(\mathbf{x}, \mathbf{x}')$ equals $n(\mathbf{x})n(\mathbf{x}')$ and thus $\Delta n(\mathbf{x}, \mathbf{x}')$ is exactly zero, i.e., the electrons are uncorrelated. In the Hartree-Fock approach, where the many electron wave function is approximated by a single Slater determinant of orthogonal orbitals, the difference $\Delta n(\mathbf{x}, \mathbf{x}')$ becomes

$$\Delta n(\mathbf{x}, \mathbf{x}') = -\delta_{ss'} \left| \sum_{i=1}^{\text{occ.}} \varphi_i^{s*}(\mathbf{r}) \varphi_i^{s}(\mathbf{r}') \right|^2 . \tag{53}$$

This demonstrates that in the Hartree-Fock approximation electrons of different spin are not correlated and behave as independent electrons. The probability of finding an electron in the

vicinity of another electron with the same spin is however reduced leading to the so-called exchange or Fermi hole. To take into account correlation effects between electrons of different spin, one has to go beyond the Hartree-Fock approximation. In particular, more than a single Slater determinant is required to render Coulomb correlation.
In terms of the total energy, correlation is usually defined as the energy difference between the exact ground-state energy and its Hartree-Fock counterpart

$$E_{\rm c} = E - E^{\rm HF} \,. \tag{54}$$

Since the Hartree-Fock total energy constitutes an upper bound for the true total energy, the correlation energy as defined in Eq. (54) must be negative. We note that this quantum chemical definition of correlation is different from that used in the field of density functional theory.
In the following, we focus on three approaches, namely MP perturbation theory, CI, and CC. All three go beyond the Hartree-Fock method and take care of Coulomb correlation.

4.1 Møller-Plesset perturbation theory

Møller-Plesset perturbation theory [3] is based on the Hartree-Fock ground-state and includes Coulomb correlation perturbatively. In fact, the interacting N electron Hamiltonian is partitioned into the N electron Fock Hamiltonian $\hat{F}$, given as a sum of the N one-particle Fock operators $\hat{f}$ and a remaining term $\hat{W}$, the perturbation,

$$\hat{H} = \hat{F} + \hat{W} \tag{55}$$

with $\hat{F} = \sum_{i=1}^{N} \hat{f}(\mathbf{x}_i)$ and $\hat{W} = \hat{H} - \hat{F}$.
The ground-state and excited states of the unperturbed system $\hat{F}$ are known: the ground-state corresponds to the Hartree-Fock single Slater determinant. Its energy is given by the expectation value $\langle \Psi^{\rm HF}|\hat{F}|\Psi^{\rm HF}\rangle$ which is equivalent to a sum over the occupied Hartree-Fock orbital energies

$$E_F^0 = \langle \Psi^{\rm HF}|\hat{F}|\Psi^{\rm HF}\rangle = \sum_{i=1}^{N} \epsilon_i \,. \tag{56}$$

Excited states are generated by replacing at least one orbital from the Hartree-Fock single Slater determinant by an unoccupied Hartree-Fock orbital. The excited states of the Fock Hamiltonian are usually categorized by their excitation level: single excitations are all Slater determinants arising from the Hartree-Fock ground-state by replacing exactly one occupied Hartree-Fock orbital by an unoccupied one. In second quantization, the singly excited Slater determinant created by substituting the occupied orbital i by the unoccupied orbital a becomes

$$|\Psi_i^a\rangle = \hat{c}_a^\dagger \hat{c}_i |\Psi^{\rm HF}\rangle \,. \tag{57}$$

The energy corresponding to the excited states $|\Psi_i^a\rangle$ is given by

$$E_F^{\ a}_{\ i} = \langle \Psi_i^a|\hat{F}|\Psi_i^a\rangle = E_F^0 + \epsilon_a - \epsilon_i \,. \tag{58}$$

Correspondingly, double excitations refer to Slater determinants and energies where two initially occupied states (i,j) are replaced by two unoccupied orbitals (a,b)

$$|\Psi_{ij}^{ab}\rangle = \hat{c}_a^\dagger \hat{c}_b^\dagger \hat{c}_i \hat{c}_j |\Psi^{\rm HF}\rangle \tag{59}$$

$$E_F^{\ ab}_{\ ij} = \langle \Psi_{ij}^{ab}|\hat{F}|\Psi_{ij}^{ab}\rangle = E_F^0 + \epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j \tag{60}$$

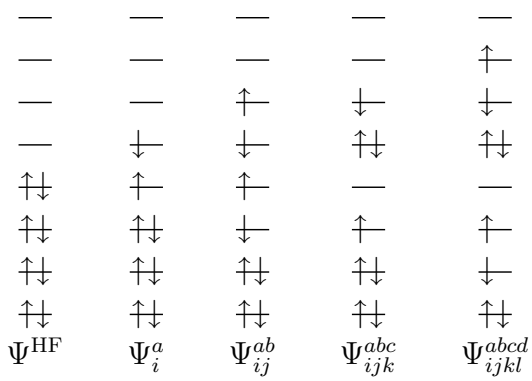


Figure 1: *Schematic representation of the Hartree-Fock ground-state reference wave function* Ψ^{HF} *and singly, doubly, triply, and quadruply excited wave functions.*

and so on and so forth (cf. Fig. 1). The highest excitation level corresponds to the case, where all electrons are promoted from an occupied to an unoccupied Hartree-Fock orbital.
In principle, the Fock operator $\hat{f}$ gives rise to an infinite number of unoccupied orbitals. In practice, however, a basis set is employed for representing the Hartree-Fock orbitals and thus the number of unoccupied states is finite and equals the difference $N_{\mathrm{basis}} - N$. The number of n-fold excited Slater determinants $N_{\mathrm{SD}}(n)$ amounts then to the possibility to choose n orbitals out of the N occupied states times the possibility to distribute them among the $N_{\mathrm{basis}} - N$ unoccupied states

$$N_{\mathrm{SD}}(n) = \left(\begin{array}{c} N \\ n \end{array} \right) \left(\begin{array}{c} N_{\mathrm{basis}} - N \\ n \end{array} \right) . \tag{61}$$

The perturbation $\hat{W}$ affects the spectrum, i.e, eigenstates as well as eigenenergies, of the Fock operator $\hat{F}$. Møller-Plesset perturbation theory calculates the change of the ground-state $|\Psi^{\mathrm{HF}}\rangle$ and its energy due to $\hat{W}$ by invoking Rayleigh-Schrödinger perturbation theory. In fact, Møller-Plesset perturbation theory is nothing more than Rayleigh-Schrödinger perturbation theory with the specific choice of the unperturbed Hamiltonian to be the Fock operator.
According to perturbation theory [16] the first order change of the ground-state energy is given by the expectation value of the perturbation with the unperturbed ground-state wave function

$$\Delta E_{\mathrm{MP}}^1 \;=\; \langle \Psi_{\mathrm{HF}} | \hat{W} | \Psi_{\mathrm{HF}} \rangle = E_{\mathrm{HF}} - E_{\mathrm{F}}^0 \, , \tag{62}$$

where we introduced the notation ΔE_{MP}^n for the n-th order change of the energy. In general, we use $\mathrm{MP}n$ as an abbreviation for n-th order MP perturbation theory. In MP1 the energy $E = E_F^0 + \Delta E_{\mathrm{MP}}^1$ is identical to the Hartree-Fock ground-state energy. At least, MP perturbation theory of second order is needed to obtain a first contribution to the correlation energy as defined in Eq. (54). However, in order to calculate the second order energy change the first order change of the wave function is required

$$|\Psi_{\mathrm{MP}}^1\rangle \;=\; \sum_{\substack{ij\ldots. \\ ab\ldots.}} \frac{\langle \Psi_{ij\ldots}^{ab\ldots} | \hat{W} | \Psi^{\mathrm{HF}} \rangle}{E_F^0 - E_F^{\substack{ij\ldots. \\ ab\ldots.}}} |\Psi_{ij\ldots}^{ab\ldots}\rangle \, . \tag{63}$$

In principle, the sum runs over all (singly, doubly, triply,...) excited Slater determinants. But, only those excited determinants contribute to the sum, for which the matrix element in the

numerator $\langle \Psi_{ij\ldots}^{ab\ldots} | \hat{W} | \Psi^{\mathrm{HF}} \rangle$ is nonzero. According to the definition of the perturbation $\hat{W} = \hat{H} - \hat{F}$ the matrix elements split up into two contributions

$$\langle \Psi_{ij\ldots}^{ab\ldots} | \hat{W} | \Psi^{\mathrm{HF}} \rangle = \langle \Psi_{ij\ldots}^{ab\ldots} | \hat{H} | \Psi^{\mathrm{HF}} \rangle - \langle \Psi_{ij\ldots}^{ab\ldots} | \hat{F} | \Psi^{\mathrm{HF}} \rangle \,, \tag{64}$$

where the second part is always zero, since both states, $|\Psi^{\mathrm{HF}}\rangle$ and $|\Psi_{ij\ldots}^{ab\ldots}\rangle$, are eigenfunctions of $\hat{F}$ corresponding to different eigenvalues and are thus orthogonal to each other. Consequently, the matrix element $\langle \Psi_{ij\ldots}^{ab\ldots} | \hat{H} | \Psi^{\mathrm{HF}} \rangle$ is decisive. As demonstrated in Appendix A matrix elements of $\hat{H}$ between the Hartree-Fock ground-state and singly excited Slater determinants vanish exactly. This is known as Brillouin theorem in the literature. Likewise, matrix elements with triply and higher excited Slater determinants vanish as the Hamiltonian $\hat{H}$ only contains two particle interactions. Thus, the only non-vanishing matrix elements are generated by doubly excited states $|\Psi_{ij}^{ab}\rangle$. By applying the rules of second quantization we obtain

$$\langle \Psi_{ij}^{ab} | \hat{H} | \Psi^{\mathrm{HF}} \rangle = \langle ab || ji \rangle (1 - \delta_{ij})(1 - \delta_{ab}) \tag{65}$$

with the abbreviation

$$\langle ab || ji \rangle = \langle ab | ji \rangle - \langle ab | ij \rangle \,. \tag{66}$$

Consequently, the first order change of the Hartree-Fock ground-state wave function is finally given by

$$|\Psi_{\mathrm{MP}}^{1}\rangle = \sum_{i}^{\mathrm{occ.}} \sum_{j<i} \sum_{a}^{\mathrm{unocc.}} \sum_{b<a} \frac{\langle ab || ji \rangle}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} |\Psi_{ij}^{ab}\rangle \,, \tag{67}$$

where we sum over all possible double excitations. With the first order change of the wave function we are in the position to calculate the second order change of the energy

$$\Delta E_{\mathrm{MP}}^{2} = \langle \Psi^{\mathrm{HF}} | \hat{W} | \Psi_{\mathrm{MP}}^{1} \rangle = \sum_{i}^{\mathrm{occ.}} \sum_{j<i} \sum_{a}^{\mathrm{unocc.}} \sum_{b<a} \frac{|\langle ab || ji \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \,. \tag{68}$$

Exploiting the symmetry of the summand in i and j and a and b and using the fact that it vanishes for $i = j$ or $a = b$, Eq. (68) can be simplified to

$$\Delta E_{\mathrm{MP}}^{2} = \langle \Psi^{\mathrm{HF}} | \hat{W} | \Psi_{\mathrm{MP}}^{1} \rangle = \frac{1}{4} \sum_{i,j}^{\mathrm{occ.}} \sum_{a,b}^{\mathrm{unocc.}} \frac{|\langle ab || ji \rangle|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} \,. \tag{69}$$

The MP2 energy involves a summation over two electron integrals between two occupied (i, j) and two unoccupied (a, b) orbitals, which in total amounts to $N^2 \times (N_{\mathrm{basis}} - N)^2$ integrals. The calculation of each integral formally scales as N_{basis}^5 leading to an overall scaling of N_{basis}^5. By using similar techniques as for speeding up Hartree-Fock calculations (integral pre-screening, resolution of the identity) [17] the scaling can be improved. Nowadays, MP2 calculations can be performed routinely for a few hundred basis functions at a cost similar to Hartree-Fock calculations.

In a similar manner, higher order energy corrections can be derived. With increasing order the

terms become more and more complicated. For example, the third order energy is given by

$$\Delta E^3_{\mathrm{MP}} = \frac{1}{8}\sum_{i,j,k,l}^{\mathrm{occ.}}\sum_{ab}^{\mathrm{unocc.}} \frac{\langle ab||kl\rangle\langle kl||ij\rangle\langle ij||ab\rangle}{(\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b)(\epsilon_k+\epsilon_l-\epsilon_a-\epsilon_b)} \tag{70}$$

$$+\frac{1}{8}\sum_{i,j}^{\mathrm{occ.}}\sum_{abcd}^{\mathrm{unocc.}} \frac{\langle ab||cd\rangle\langle cd||ij\rangle\langle ij||ab\rangle}{(\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b)(\epsilon_i+\epsilon_j-\epsilon_c-\epsilon_d)} \tag{71}$$

$$+\sum_{i,j,k}^{\mathrm{occ.}}\sum_{abc}^{\mathrm{unocc.}} \frac{\langle ij||ab\rangle\langle kb||cj\rangle\langle ac||ik\rangle}{(\epsilon_i-\epsilon_j-\epsilon_a-\epsilon_b)(\epsilon_i-\epsilon_k-\epsilon_a-\epsilon_c)}\,. \tag{72}$$

The evaluation of the MP3 energy term formally requires N^6_{basis} operations. In comparison to MP2, it scales worse by a factor of N_{basis}.
Generally, the computational complexity and cost increases with the order of the perturbation. The calculation of the n-th MP energy (MPn) (and $n \geq 2$) formally scales with N^{3+n}_{basis} making calculations for n larger than 4 computationally very demanding. In order to demonstrate the consequences of the scaling behavior, we assume that a MP2 (MP3) calculation for a given system and a given basis requires 6 hours. Then, after doubling of the basis-set size to check the convergence of the calculation it lasts $8\,(16)$ days.
By comparing to (full) CI calculations (explained and discussed in the next section) MP2 typically accounts for $80-90\%$ of the correlation energy. MP3 usually covers $90-95\%$ of the energy and with MP4, which exhibits an N^7_{basis} scaling, up to 98% are considered. This demonstrates, that on the one hand side with increasing order n a larger amount of the correlation energy is reproduced. On the other side, however, the (percental) gain in the correlation energy with increasing n becomes smaller and smaller. In other words, one has to work harder and harder to account for the remaining few percents of the correlation energy.
In contrast to Hartree-Fock, MP perturbation theory is not variational. Hence, it can happen that the MPn total energy is smaller than the exact total energy. Moreover, no systematic convergence of the MP energy is guaranteed [18, 19]. In the ideal case, the convergence is monotonic. However, depending on the system also an oscillating convergence behavior can be observed. In the worst case, if the fundamental assumption of perturbation theory, that the perturbation is small, is not valid, the MPn energies may even diverge.
In general, MP perturbation theory usually constitutes a good compromise between accuracy and computational cost.

4.2 Configuration Interaction method

In contrast to Møller-Plesset perturbation theory, configuration interaction (CI) does not rely on perturbation theory. CI, instead, is based on the fact that the space spanned by all N electron Slater determinants formed from a complete set of one-particle spin-orbitals constitutes an exact basis for expanding the antisymmetric N electron wave function. Hence, starting from a so-called reference wave function $|\Psi_0\rangle$, for which typically the Hartree-Fock ground-state wave function is employed, the exact many electron wave function can be expanded into a linear combination of $|\Psi_0\rangle$, single, double, and higher excitations of $|\Psi_0\rangle$

$$|\Psi\rangle = |\Psi_0\rangle + \sum_{i,a} C^a_i|\Psi^a_i\rangle + \sum_{\substack{i,j\\ i<j}}\sum_{\substack{a,b\\ a<b}} C^{ab}_{ij}|\Psi^{ab}_{ij}\rangle + \dots\,. \tag{73}$$

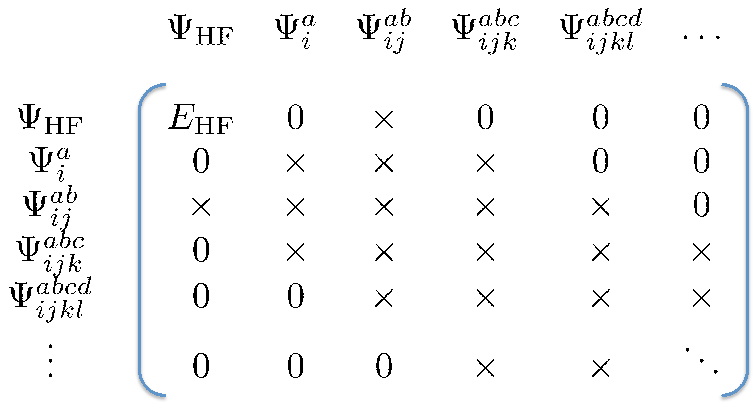


Figure 2: *Structure of the CI Hamiltonian: non-zero elements are marked with a cross and each excited Slater determinant is in fact a representative for the whole set of n-fold excited determinants.*

The set of singly, doubly, and higher excited Slater determinants are constructed from $|\Psi_0\rangle$ according to Eqs. (57) and (59). The only remaining unknowns in Eq. (73) are the expansion coefficients $\{C_i^a, C_{ij}^{ab}, \dots\}$. These are determined by invoking the variational principle, i.e., the expectation value $\langle\Psi|\hat{H}|\Psi\rangle$ is minimized with respect to $\{C_i^a, C_{ij}^{ab}, \dots\}$ subjected to the constraint that $|\Psi\rangle$ is normalized. The minimization procedure transfers the variational problem into an algebraic one

$$(\mathbf{H} - E\mathbf{I})\,\mathbf{C} = 0\,, \tag{74}$$

where we subsumed the expansion coefficients $\{C_i^a, C_{ij}^{ab}, ...\}$ into the vector $\mathbf{C}$ and ordered the different Slater determinants accordingly. The matrix elements of $\mathbf{H}$ are then given by

$$H_{ij} = \langle\Psi_i|\hat{H}|\Psi_j\rangle \tag{75}$$

and $\mathbf{I}$ denotes the unit matrix. (For simplicity, we assumed that the spin orbitals forming the Slater determinants are orthonormal.) Equation (74) corresponds to an eigenvalue problem, whose lowest eigenvalue and eigenvector corresponds to the CI ground-state energy and the CI ground-state wave function, respectively. We note that the eigenvalue problem must be solved only once: it does not constitute a self-consistent field problem.

Hamiltonian matrix elements between Slater determinants which differ by more than two orbitals are zero since the Hamiltonian consists only of one- and two-electron operators. Moreover, due to the Brillouin theorem (s. Appendix A) matrix elements of a singly excited Slater determinant with the Hartree-Fock ground-state are zero, as well. Hence, the CI Hamiltonian acquires a band matrix structure as shown in Fig. 2.

While the procedure so far is exact under the assumption of a complete (and thus infinite) basis for the one-particle spin-orbitals, the usage of a finite (and thus incomplete) basis is inevitable in practice. With the consequence that the expansion, Eq. (73), is fulfilled only approximately. But even a finite one-particle basis with N_{basis} basis functions can give rise to a gigantic number N_{SD} of Slater determinants

$$N_{\mathrm{SD}} = \sum_{n=0}^{N} N_{\mathrm{SD}}(n) = \sum_{n=0}^{N} \left(\begin{array}{c} N \\ n \end{array} \right) \left(\begin{array}{c} (N_{\mathrm{basis}} - N) \\ n \end{array} \right) = \left(\begin{array}{c} N_{\mathrm{basis}} \\ N \end{array} \right) \tag{76}$$

Basis set	N_{basis}	N_{SD}
cc-pVDZ	14	1001
cc-pVTZ	30	30045015
cc-pVQZ	55	29248649430

Table 1: *Number of Slater determinants for the Ne atom ($N = 10$) and three different basis sets.*

where N denotes the number of electrons and n is the excitation level. In fact, the number of Slater determinants N_{SD} grows factorial with the size of the basis. As an example we show in Table 1 the total number of Slater determinants for the Ne atom with 10 electrons $(N = 10)$ and for three different basis sets of increasing quality. Already the triple zeta basis set (cc-pVTZ) [20, 6] gives rise to a tremendous number of Slater determinants.[4] The diagonalization of the corresponding Hamiltonian matrix presents a formidable computational task even for the present generation of supercomputers. Consequently, CI calculations taking into account all possible (excited) Slater determinants of a given basis set, so-called full CI (FCI) calculations, are only possible for the smallest systems with a small basis set.
To reduce the computational workload of (full) CI calculations and to enable calculations for larger systems, one goes over to truncated CI. In a truncated CI calculation only specific classes of excited Slater determinants are considered in the expansion of the wave function [Eq. (73)]. For example, CI single (CIS) truncates the expansion after the set of single excitations. Since single excitations do not couple to the HF ground-state wave function (s. Fig. 2), a CIS calculation does not improve on the Hartree-Fock ground-state energy, i.e., the lowest eigenvalue of the CIS Hamiltonian is identical to the HF ground-state energy. Though, the higher eigenvalues of the CIS matrix can be used as approximate energies for singly excited electronic states. In the CI single double (CISD) approach singly and doubly excited Slater determinants are used to expand the many-electron wave function and an improved ground-state energy is obtained. Accordingly higher truncated CI schemes are defined.
The computational complexity of truncated CI methods relies on the truncation level. While a CISD calculation scales as N_{basis}^6, CISDT exhibits a N_{basis}^8 scaling and the next higher order (CISDTQ) shows a N_{basis}^{10} behavior [21]. Only the CISD approach is generally applicable to a large variety of systems. Typically CISD covers about 90% of the correlation energy.
While the truncation of the CI approach is absolutely necessary to restrict the number of Slater determinants and thus the size of the CI Hamiltonian to a feasible value, the truncation, however, destroys two important properties of full CI namely size consistency and size extensivity. An approach is called size consistent if the energy of a system formed by two *non-interacting* subsystems A and B equals the energy of the individual system A plus the energy of B, i.e.,

$$E_{\mathrm{AB}} = E_{\mathrm{A}} + E_{\mathrm{B}}\,. \tag{77}$$

Size consistency is decisive for describing the dissociation of molecules. In the dissociation limit, corresponding to an infinite distance between the constituting atoms of the molecule, size consistency guarantees the correct limit of the dissociation energy. On the other hand, size extensivity deals with the correct scaling of the (correlation) energy with the system size. In

[4] The number of Slater determinants can be reduced by considering only those Slater determinants of a given spatial and spin symmetry.

fact, an approach is called size extensive, if the (correlation) energy scales in the limit of a large system with the system size

$$\lim_{N\to\infty} \frac{E(N \times A)}{N} = \text{const.} \,. \tag{78}$$

Size extensivity ensures that the energy does not diverge with increasing system size and is of utmost importance for dealing with infinite systems like crystalline solids. Note that both properties are fulfilled in a full CI calculation. Truncation of the CI expansion [Eq. 73] at a finite order, however, causes the violation of size consistency and size extensivity. As a consequence, the quality of truncated CI calculations decreases with increasing system size. Both drawbacks are cured by the Coupled Cluster approach, which we discuss in the next section.

4.3 Coupled Cluster method

The Coupled Cluster (CC) method is nowadays the method of choice for high accuracy calculations (if computational feasible). In analogy to the CI approach, the many-electron wave function is written in terms of an excitation operator acting on a single reference Slater determinant, typically the Hartree-Fock ground-state. However, while CI uses a linear excitation operator, an exponential form is invoked in the CC method

$$|\Psi^{\text{CI}}\rangle = (1+\hat{T})|\Psi_0\rangle = (1+\hat{T}_1+\hat{T}_2+\hat{T}_3+\dots)|\Psi_0\rangle \tag{79}$$

$$|\Psi^{\text{CC}}\rangle = \exp(\hat{T})|\Psi_0\rangle = \exp(\hat{T}_1+\hat{T}_2+\hat{T}_3+\dots)|\Psi_0\rangle \,, \tag{80}$$

where for example the single excitation operator $\hat{T}_1$ is defined by $\hat{T}_1 = \sum_i^{\text{occ.}} \sum_a^{\text{unocc.}} C_i^a \, \hat{c}_a^\dagger \hat{c}_i$ [cf. with Eq. (73)]. Double and higher-order excitations operators are defined analogously. Without truncation of the excitation operator, both approaches, CI and CC, are exact (in the complete basis set limit) and give rise to identical results. Yet, truncation is a prerequisite in practice. The exponential form of the truncated excitation operator has the advantage, that despite the truncation of $\hat{T}$ at a finite order n, higher orders are included through the exponential. For example, in a CCD calculation, which approximates $\hat{T}$ simply by $\hat{T}_2$, the CC wave function is due to the exponential ansatz a linear combination of doubly, quadruply, hextubly up to N-tuply excited Slater determinants (where N denotes again the number of electrons). However, the number of independent coefficients in comparison to a CID calculation are exactly the same, since the coefficients for quadruble and higher excitations are products of lower order coefficients. In the CC method, the so-far unknown coefficients $\{C_i^a, C_{ij}^{ab}, \dots\}$ are typically called cluster amplitudes.

The CC amplitudes and total energy are obtained by inserting the CC wave function in the Schrödinger equation

$$\hat{H} \, \exp(\hat{T})|\Psi_0\rangle = E \, \exp(\hat{T})|\Psi_0\rangle \,. \tag{81}$$

By multiplying this equation with the bra $\langle\Psi_0|$ and exploiting the orthogonality of an excited Slater determinant with respect to Ψ_0, the total energy E may be obtained from

$$E = \langle\Psi_0|\hat{H} \exp(\hat{T})|\Psi_0\rangle \,. \tag{82}$$

Since the Hamiltonian $\hat{H}$ only contains one- and two-electron operators, the in principle infinite Taylor expansion of the exponential operator terminates at the double excitation level

$$E = \langle\Psi_0|\hat{H}(1+\hat{T}_1+\hat{T}_2+\frac{1}{2}\hat{T}_1^2)|\Psi_0\rangle \,. \tag{83}$$

Hence, the CC total energy is fully determined by the single and double cluster amplitudes. This, however, does not imply that a CCSD calculation is already sufficient to obtain the correct total energy. As we will discuss below, single and double cluster amplitudes are coupled to the amplitudes of higher excitations. Consequently, the total energy implicitly depends on the higher cluster amplitudes, as well.
Equations for the cluster amplitudes arise from the projection of the Schrödinger equation [Eq. (81)] on the excited Slater determinants

$$\begin{aligned} \langle \Psi_i^a | \hat{H} \exp(\hat{T}) | \Psi_0 \rangle &= E \langle \Psi_i^a | \exp(\hat{T}) | \Psi_0 \rangle & (84) \\ \langle \Psi_{ij}^{ab} | \hat{H} \exp(\hat{T}) | \Psi_0 \rangle &= E \langle \Psi_{ij}^{ab} | \exp(\hat{T}) | \Psi_0 \rangle & (85) \\ &\vdots \quad . \end{aligned}$$

This set of equations for the cluster amplitudes is coupled with the equation for the total energy. It is more convenient and suitable in practice to separate the equations for the cluster amplitudes from that for the total energy. This can be achieved by multiplying the Schrödinger equation [Eq. (81)] first with $\exp(-\hat{T})$ and then projecting onto the excited determinants

$$\begin{aligned} \langle \Psi_i^a | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \Psi_0 \rangle &= 0 & (86) \\ \langle \Psi_{ij}^{ab} | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \Psi_0 \rangle &= 0 & (87) \\ &\vdots \quad . \end{aligned}$$

As an example, we demonstrate that the first equation [Eq. (86)] gives rise to an algebraic equation, which couples single, double and triple amplitudes. By acting with the adjoint operator of $\exp(-\hat{T})$ to the left, Eq. (86) becomes

$$\langle \Psi_i^a | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) | \Psi_0 \rangle = \langle \exp(-\hat{T}^\dagger) \Psi_i^a | \hat{H} | \exp(\hat{T}) \Psi_0 \rangle \, . \tag{88}$$

The adjoint excitation operator $\hat{T}^\dagger$ functions as a de-excitation operator, e.g.,

$$\hat{T}_1^\dagger = \sum_i^{\text{occ.}} \sum_a^{\text{unocc.}} C_i^{a*} (\hat{c}_a^\dagger \hat{c}_i)^\dagger = \sum_i^{\text{occ.}} \sum_a^{\text{unocc.}} C_i^{a*} \hat{c}_i^\dagger \hat{c}_a \, . \tag{89}$$

Hence, $\exp(-\hat{T}^\dagger)\Psi_i^a$ generates in addition to the singly excited state the reference wave function. Taking into account the nature of the Hamilton operator we end-up with the following equation

$$\langle (1 - \hat{T}_1^\dagger) \Psi_i^a | \hat{H} | [1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2}\hat{T}_1^2) + (\hat{T}_3 + \hat{T}_2\hat{T}_1 + \frac{1}{6}\hat{T}_1^3)] \Psi_0 \rangle = 0 \, , \tag{90}$$

in which the single, double, and triple amplitudes enter through the corresponding (de-)excitation operators as variables. In a similar manner, further equations for the cluster amplitudes arise by projection onto the higher excited Slater determinants. The so-obtained set of coupled, non-linear equations, which is of dimension $N \times N$ and N equals the number of cluster amplitudes, must be solved. Usually, an iterative approach, e.g., a quasi-Newton approach, is employed for solving this set of non-linear equations. The so-obtained cluster amplitudes are then used to calculate the total energy according to Eq. (82).
It is instructive to plug in the order-by-order iterative solution for the cluster amplitudes in the CC energy equation [Eq. (82)]. In this way terms of Møller-Plesset perturbation theory are

recovered as demonstrated in Ref. [8]. This shows that CC theory can be interpreted as an infinite resummation of selected terms of perturbation theory.
The CC single (CCS) approach similar to CIS does not improve on the Hartree-Fock ground-state energy. In fact, it yields the Hartree-Fock energy. At least double excitations are necessary to render electron correlation. The CCD as well as CCSD method exhibit an formal scaling of N_{basis}^6. Taking into account triple excitations in addition, the scaling becomes proportional to N_{basis}^8 restricting the CCSDT approach to small systems with a small number of basis functions. The exponential ansatz for the excitation operator and thus the consideration of an excitation type up to infinite order makes CC theory in contrast to CI an size extensive and size consistent approach.[5] In combination with the fact that truncated CI approaches show the same formal scaling as the corresponding CC methods, truncated CC is nowadays the method of choice for high accuracy calculations of the ground-state wave function and energy.
We note that we discussed the standard formulation of CC theory, which is not variational. A variational formulation of CC theory is possible in principle [10]. However, variational coupled cluster is more demanding than the presented standard formulation restricting its range of application to the smallest systems. Moreover, excited states are not accessible in the presented ground-state formalism. Extensions of the CC formalism, as for example equation-of-motion coupled cluster (EOM-CC) or linear-response coupled cluster (LR-CC) [22], enable the calculation of excited states.

5 Application to the diatomic molecule BH

In this section the aforementioned approaches are applied to the diatomic molecule BH, which was first observed experimentally by Lochte-Holtgreven and van der Vleugel in 1931 [23]. Since then it has been intensively studied experimentally and theoretically. Its small size makes BH an ideal candidate for testing and validating electronic structure methods. In fact, the first ab-initio study of BH goes back to Shani in 1956 [24].
Table 2 shows the correlation energy obtained for the BH molecule at different levels of theory, comprising Møller-Plesset perturbation theory, CI, and CC. In addition, the result of a full CI calculation, corresponding to the exact result for the given basis set, is presented. The percental values are given with respect to the full CI correlation energy. All calculations are performed with a polarized double-zeta basis set consisting of 19 basis functions. The lowest order of MP perturbation theory that accounts for correlation (MP2) covers 67.37% of the total correlation energy of BH. By increasing the order n of perturbation theory a larger and larger amount of the correlation energy is accounted for. We note that BH belongs to the (lucky) case where MP perturbation theory exhibits a monotonic convergence from above to the full CI result (at least up to the 7-th order). In general, it is not guaranteed that MP perturbation theory shows a convergent behavior. Moreover, since it is not a variational approach, it does not have to reach the exact result from above. The truncated CI calculations are opposed to those MPn calculations, which exhibit the same formal scaling. Hence, we compare CISD with MP3, CISDT with MP5, and CISDTQ with MP7. The CISD calculation already accounts for 95% of the whole correlation energy, which is a significantly larger portion than covered by MP3. By going over to CISDT or even CISDTQ the error in the correlation energy is reduced to $3.04\,\%$ and $0.3‰$, respectively. In contrast to perturbation theory, CI is variational. Thus, at each truncation level the CI result poses an upper bound for the true correlation energy. We note that

[5] Size consistency demands that the reference wave function fulfills this property [10].

BH	$E_\mathrm{c}\,(\mathrm{mHtr})$			$E_\mathrm{c}\,(\mathrm{mHtr})$			$E_\mathrm{c}\,(m\mathrm{Htr})$	
MP*n*	abs.	%	CI*n*	abs.	%	CC*n*	abs.	%
MP2	-60.723	67.37						
MP3	-78.558	87.15	CISD	-85.670	95.04	CCSD	-88.284	97.94
MP4	-84.910	94.20						
MP5	-87.565	97.15	CISDT	-87.398	96.96	CCSDT	-90.070	99.93
MP6	-88.821	98.54						
MP7	-89.460	99.25	CISDTQ	-90.107	99.97	CCSDTQ	-90.137	100.00
FCI	-90.137							

Table 2: *Correlation energy for the BH molecule using Møller-Plesset perturbation, CI, and CC theory and a polarized double-zeta basis set consisting of 19 basis functions. The percental value is calculated with respect to the full CI (FCI) correlation energy. The Hartree-Fock reference energy amounts to* $25.125187\,\mathrm{Htr}$.
(The data is taken from Ref. [19] apart from the CISDTQ and CCSDTQ results. These are from Refs. [25] and [26]. We note that slightly different bond distances are employed in the latter calculations.)

truncated CI exhibits diminished accuracy for larger systems as it is not size-extensive. Finally, we turn to the Coupled Cluster calculations for BH. At each level of truncation the CC method accounts for the largest amount of the full correlation energy compared to corresponding CI and MPn calculations. Typically, the CC correlation energy converges from above to the FCI result. Yet, this is not ensured as long as a non-variational formulation of CC theory is employed.

So far, all results have been obtained with a fixed basis set. However, the basis set is an important convergence parameter. This is not surprising as the basis set size determines on the one hand the number of unoccupied states that enter in each order of MP perturbation theory and on the other hand it defines the number of excited Slater determinants of CI and CC theory. For the case of BH we demonstrate in Table 3 that the absolute value of the correlation energy exhibits a sizeable gain with increasing basis set size. In fact, the changes caused by the basis are of the same size or even larger than those arising from different levels of theory (as discussed in Table 2). A thorough convergence of the basis is indispensable for reliable and accurate results. It is a matter of experience that basis set convergence becomes more and more important and at the same time more and more cumbersome for systems where correlation effects play a dominant role. We note by passing that extrapolation schemes have been developed to reach the basis-set limit [27].

Finally, we show results for the calculated equilibrium bond distance between the B and H atom in the ground-state of the BH molecule. Table 4 summarizes the obtained bond distance r_e for various levels of theory and three different basis sets. We first observe that the bond length consistently decreases with increasing basis set size. Hence, a small basis set tends to an overestimation of the theoretical bond length. Furthermore, it becomes evident that taking into account electron correlation by a more and more sophisticated approach leads to an enhancement of r_e. It is a general fact that Hartree-Fock theory (in the basis set limit) underestimates the bond length. Since electron-electron correlation is only treated approximately in the Hartree-Fock approach the electrons tend to bunch together and form too compact bonds. Taking into

BH	$E_\mathrm{c}\,(\mathrm{mHtr})$		
	cc-pVDZ	cc-pVTZ	cc-pVQZ
MP3	-78.558	-90.087	-93.263
MP5	-87.565	-98.612	-101.641
MP7	-89.460	-100.483	-103.516
CISD	-85.670	-95.743	-98.571
CISDT	-87.398	-98.102	-101.074
CCSD	-88.284	-98.674	-101.562
CCSDT	-90.070	-101.145	-104.193
FCI	-90.137	-101.231	-104.282
HF	-25125.187	-25129.905	-25131.286

Table 3: *Convergence of the correlation energy of the BH molecule with respect to the basis set. Three different basis sets (cc-pVDZ, cc-pVTZ, and cc-pVQZ) of increasing size and quality are employed. The cc-pVDZ basis set consists of 19 basis functions. The cc-pVTZ basis comprises 44 functions and 85 functions are used in the cc-pVQZ basis. For reference the Hartree-Fock total energy is given, as well. (The data has been taken from Ref. [19].)*

account electron correlation by a higher level of theory moves the electrons further apart and the bond length increases. In conclusion, there is an opposing trend of correlation and basis set size, which tends to increase and decrease bond distances, respectively. In order to obtain reliable molecular structures it is necessary to use a properly converged basis sets with a high-level approach for electron correlation. In fact, the CCSDT and FCI result with the cc-pVQZ basis set overestimates the experimental bond distance by about $1.06\,‰$. In the limit of a complete basis an even better agreement is to be expected. However, on this accuracy level other error sources as relativistic effects or corrections due to the Born-Oppenheimer approximation become decisive. The influence of both corrections has been studied and quantified in Ref. [28].

The vast majority of calculations employ MP perturbation theory, CI, or CC to atomic and molecular systems. The application of these approaches to solids, i.e., infinite periodic systems, has been prohibited for a long time due to the involved enormous computational demand. Only recently MP2 and CCSD calculations for solids using a plane-wave basis have been reported in the literature [30, 31, 32]. Note that the CI method is not suited for solids since it is not size extensive.

6 Summary

The Hartree-Fock method is the basis for a variety of wave-function based quantum chemical approaches that explicitly take into account electron-electron correlation. While electron correlation is considered in a perturbative manner on top of the Hartree-Fock Hamiltonian in Møller-Plesset perturbation theory, the CI and CC method formulate the many-electron wave function in terms of an excitation operator acting on a reference wave function, for which usually the Hartree-Fock wave function is employed. A systematic convergence towards the exact

BH	r_e (Å)		
	cc-pVDZ	cc-pVTZ	cc-pVQZ
HF	1.23596	1.22206	1.22030
MP3	1.25186	1.23130	1.22928
MP5	1.25646	1.23605	1.23396
MP7	1.25692	1.23648	1.23438
CISD	1.25345	1.23322	1.23114
CISDT	1.25459	1.23437	1.23234
CCSD	1.25480	1.23435	1.23219
CCSDT	1.25603	1.23557	1.23346
FCI	1.25597	1.23560	1.23349
Expt.	1.23218		

Table 4: *Calculated and experimental equilibrium bond distance r_e of the BH molecule in its ground-state. The calculated bond distance is shown for different quantum chemical approaches and different basis sets. (The theoretical values are from Ref. [19], whereas the experimental bond distance is taken from Ref. [29].)*

solution is in principle possible for all three approaches. The accuracy of a calculation depends on the quality of the employed basis set and the level of theory. In the limit of a complete basis, the different levels of theory can be typically ordered in the following way

$$\mathrm{HF} \ll \mathrm{MP2} < \mathrm{CISD} \approx \mathrm{MP4} \approx \mathrm{CCSD} < \mathrm{CCSDT} < \mathrm{full\,CI} = \mathrm{full\,CC}\ . \tag{91}$$

The price for the accuracy gain along the series is an enormous increase in computational workload. While the Hartree-Fock approach formally exhibits a N_basis^4 scaling, CISD and CCSD scales with the sixth power of the basis set size (N_basis^6). The next higher order, i.e., CISDT and CCSDT calculations, show an N_basis^8 behavior. Full CI (full CC) calculations are only feasible for small systems using only a few basis functions. We close by noting that with increasing accuracy relativistic effects and the Born-Oppenheimer approximation might become important.

Acknowledgements

The author thanks Iris Theophilou, Timo Schena and Mathias C. T. D. Müller for a critical reading of the manuscript.

Appendices

A Brillouin theorem

The Brillouin theorem states that the matrix element of the Hamiltonian $\hat{H}$ between the Hartree-Fock ground-state wave function Ψ^{HF} and a singly excited wave function Ψ_a^i vanishes

$$\langle \Psi_i^a | \hat{H} | \Psi^{\mathrm{HF}} \rangle = 0 \,. \tag{92}$$

We prove this theorem by using the Hamiltonian in second quantization [Eq. (9)]. Then, the matrix element splits up into two contributions

$$\sum_{pq} h_{pq} \langle \Psi_i^a | \hat{c}_p^\dagger \hat{c}_q | \Psi^{\mathrm{HF}} \rangle + \frac{1}{2} \sum_{pqrs} \langle pq|rs \rangle \langle \Psi_i^a | \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_s \hat{c}_r | \Psi^{\mathrm{HF}} \rangle \,. \tag{93}$$

Exploiting the orthogonality of different Slater determinants the first part which arises from the one-electron Hamiltonian is nonzero only if

$$\langle \Psi_i^a | \hat{c}_p^\dagger \hat{c}_q | \Psi^{\mathrm{HF}} \rangle = \delta_{pa} \delta_{qi} \tag{94}$$

holds. With the definition Eq. (57) the matrix element comprising the two-electron part of the Hamiltonian turns into

$$\langle \Psi_i^a | \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_s \hat{c}_r | \Psi^{\mathrm{HF}} \rangle = \langle \hat{c}_p \hat{c}_a^\dagger \hat{c}_i \Psi^{\mathrm{HF}} | \hat{c}_q^\dagger \hat{c}_s \hat{c}_r \Psi^{\mathrm{HF}} \rangle = \langle (\delta_{pa} - \hat{c}_a^\dagger \hat{c}_p) \hat{c}_i \Psi^{\mathrm{HF}} | \hat{c}_q^\dagger \hat{c}_s \hat{c}_r \Psi^{\mathrm{HF}} \rangle \,, \tag{95}$$

where the commutation relation Eq. (16) has been used. Making again use of the orthogonality condition we obtain

$$\begin{aligned} \langle \Psi_i^a | \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_s \hat{c}_r | \Psi^{\mathrm{HF}} \rangle &= \delta_{pa} \langle \hat{c}_i \Psi^{\mathrm{HF}} | \hat{c}_q^\dagger \hat{c}_s \hat{c}_r \Psi^{\mathrm{HF}} \rangle - \langle \hat{c}_a^\dagger \hat{c}_p \hat{c}_i \Psi^{\mathrm{HF}} | \hat{c}_q^\dagger \hat{c}_s \hat{c}_r \Psi^{\mathrm{HF}} \rangle & (96) \\ &= n_r n_s (1 - \delta_{sr}) \left\{ \delta_{pa} [\delta_{ir} \delta_{qs} - \delta_{is} \delta_{qr}] - \delta_{qa} [\delta_{ps} \delta_{ir} - \delta_{pr} \delta_{is}] \right\} \,. & (97) \end{aligned}$$

The occupation numbers n_r and n_s, that are one if the state is occupied and zero otherwise, emerge from the fact that the annihilation operator $\hat{c}_r$ acting on the Hartree-Fock ground-state wave function Ψ^{HF} yields zero if the one-particle orbital r is unoccupied. The term $(1 - \delta_{sr})$ excludes the case that the indices s and r are identical, which corresponds to the successive application of the same annihilation operator and gives rise to zero.
Combining Eqs. (94), (97), and (93) then yields

$$\langle \Psi_i^a | \hat{H} | \Psi^{\mathrm{HF}} \rangle = h_{ai} + \sum_r^{\mathrm{occ.}} (1 - \delta_{ri}) (\langle ar|ir \rangle - \langle ar|ri \rangle) \,, \tag{98}$$

which is identical to off-diagonal matrix element of the Fock Hamiltonian $\hat{f}$ between the unoccupied a and occupied i Fock orbital

$$\langle \Psi_i^a | \hat{H} | \Psi^{\mathrm{HF}} \rangle = \langle \varphi_a | \hat{f} | \varphi_i \rangle = \epsilon_i \langle \varphi_a | \varphi_i \rangle \,. \tag{99}$$

From the orthogonality of the Fock orbitals φ_a and φ_i the Brillouin theorem, Eq. (92), finally results.

References

[1] M. Born and J. R. Oppenheimer, Ann. der Physik **389**, 457 (1927).

[2] V. Fock, Z. Physik **61**, 126 (1930).

[3] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[4] I. Shavitt, Molecular Physics **94**, 3 (1998).

[5] J. Cizek, J. Chem. Phys. **45**, 4256 (1966).

[6] F. Jensen, *Introduction to Computational Chemistry* (Wiley, 2007).

[7] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (McGraw-Hill, 1989).

[8] R. J. Bartlett and J. F. Stanton, Reviews in Computational Chemistry **5**, 65 (1994).

[9] C. D. Sherrill and H. F. Schaefer III, Adv. Quant. Chem. **34**, 143 (1999).

[10] T. D. Crawford and H. F. Schaefer III, Reviews in Computational Chemistry **14**, 33 (2000).

[11] R. J. Bartlett and M. Musial, Rev. Mod. Phys. **79**, 291 (2007).

[12] R. Pauncz, *Spin Eigenfunctions: Construction and Use* (Plenum, 1979).

[13] T. Koopmans, Physica **1**, 104 (1934).

[14] M. Häser and R. Ahlrichs, Journal of Computational Chemistry **10**, 104 (1989).

[15] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, Chem. Phys. Lett. **240**, 283 (1995).

[16] E. Schrödinger, Ann. Phys. **385**, 437 (1926).

[17] M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).

[18] T. H. Dunning and K. A. Peterson, J. Chem. Phys. **108**, 4761 (1998).

[19] M. L. Leininger, W. D. Allen, H. F. Schaefer III, and C. D. Sherill, J. Chem. Phys. **112**, 9213 (2000).

[20] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[21] C. D. Sherrill (1996), http://vergil.chemistry.gatech.edu/notes/ciscale/ciscale.pdf (last access 2013-12-03).

[22] J. D. Watts, Challenges and Advances in Computational Chemistry and Physics **5**, 65 (2008).

[23] W. Lochte-Holtgreven and E. S. van der Vleugel, Z. Phys. **70**, 188 (1931).

[24] R. C. Shani, J. Chem. Phys. **25**, 332 (1956).

[25] R. J. Harrison and N. C. Handy, Chem. Phys. Lett. **95**, 386 (1983).

[26] S. A. Kucharskia and R. J. Bartlett, J. Chem. Phys. **97**, 4282 (1992).

[27] D. G. Truhlar, Chemical Physics Letters **294**, 45 (1998).

[28] B. Temelso, E. F. Vallev, and C. D. Sherrill, J. Chem. Phys. A **108**, 3068 (2004).

[29] F. S. Pianalto, L. C. O'Brien, P. C. Keller, and P. F. Bernath, J. Mol. Spectrosc. **129**, 348 (1988).

[30] M. Marsman, A. Grüneis, J. Paier, and G. Kresse, J. Chem. Phys. **130**, 184103 (2009).

[31] A. Grüneis, M. Marsman, and G. Kresse, J. Chem. Phys. **133**, 074107 (2010).

[32] A. Grüneis, G. H. Booth, M. Marsman, J. Spencer, A. Alavi, and G. Kresse, J. Chem. Theory Comput. **7**, 2780 (2011).

A 4 Many-Body Perturbation Theory: The *GW* Approximation [1]

Christoph Friedrich and Arno Schindlmayr
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

In the previous lectures we have seen that density-functional theory (DFT) is the method of choice when we are interested in the ground-state properties of a many-electron system. DFT is based on the Hohenberg-Kohn theorem [1], which states that there is (a) a one-to-one correspondence between the ground-state density $n_0(\mathbf{r})$ and the external potential as well as (b) a variational principle for the energy functional $E[n_0] \leq E[n]$. The second statement allows to obtain the ground state of a many-electron system by variation of its density, a quantity that is much less complicated than the many-electron wave function $\Psi_0(\mathbf{r}_1, ..., \mathbf{r}_N)$, where N is the particle number. The first statement implies that the many-particle Hamiltonian is a functional of the ground-state density. Since the diagonalization of the Hamiltonian yields the complete excitation spectrum, the excited states can ultimately be regarded as functionals of the ground-state density as well. However, the Hohenberg-Kohn theorem does not provide us with an explicit mathematical form. In this lecture we show that excited-state properties can be accessed more directly with a purpose-built method, the so-called many-body perturbation theory [2, 3]. Incidentally, in practice its implementation within the GW approximation [4] for the electronic self-energy is based on a perturbative evaluation with Kohn-Sham orbitals and can, therefore, finally be interpreted as the desired density functional.

The solution of the Kohn-Sham equation [5] of DFT yields a whole spectrum of single-particle states, and one is tempted to identify the corresponding eigenvalues with excitation energies. Strictly speaking, such an interpretation is wrong: the Kohn-Sham wave functions and eigen-

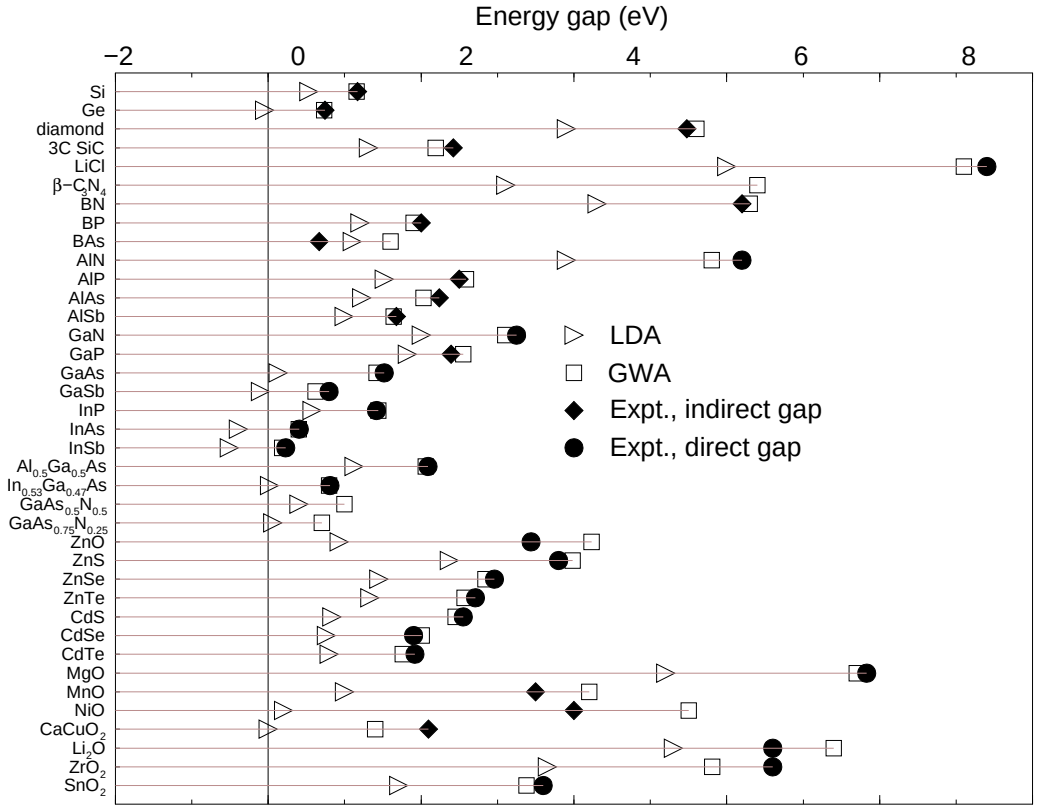


Fig. 1: *Comparison of LDA, GW and experimental band gaps for a variety of materials. Taken from Ref. [8].*

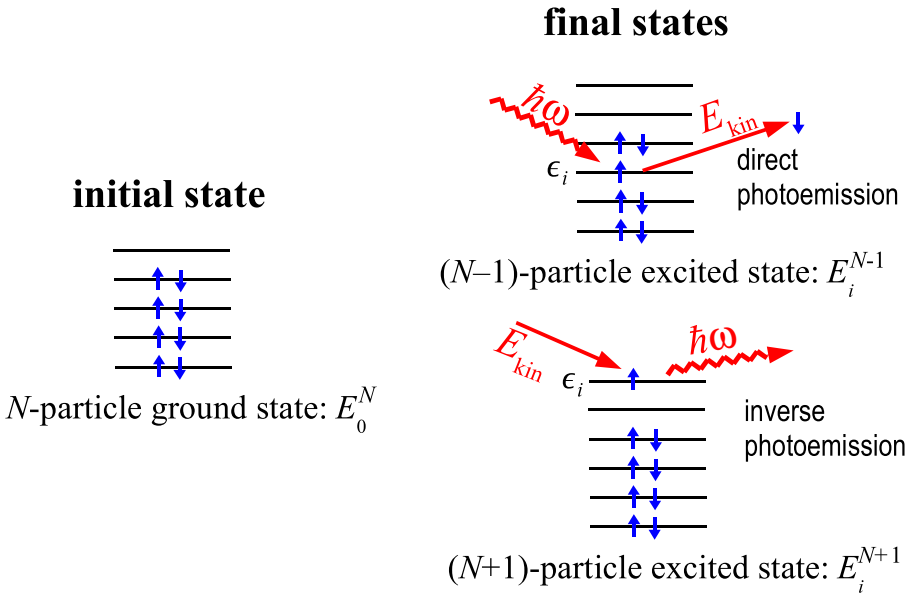


Fig. 2: *Schematic illustration of direct and inverse photoelectron spectroscopy. In both processes the particle number changes. The measured energy difference* $E_{\rm kin} - \hbar\omega$ *corresponds to* $\epsilon_i = E_0^N - E_i^{N-1}$ *in direct and* $\epsilon_i = E_i^{N+1} - E_0^N$ *in inverse photoelectron spectroscopy.*

values must be considered as mathematical tools and cannot be endowed with a physical meaning. The only exception is the energy of the highest occupied state, which equals the exact ionization potential (or chemical potential for metals) [6, 7]. Consequently, while often qualitatively correct, the DFT band structure fails to give reliable quantitative values for the band gaps of insulators and semiconductors, which are often underestimated by as much as 1.0 eV or more. In the case of Ge the local-density approximation (LDA) of DFT even predicts a semi-metal with a negative band gap rather than a semiconductor. In this lecture we demonstrate that the Kohn-Sham eigenvalues can be corrected using Green-function techniques and the GW approximation for the electronic self-energy. Figure 1 shows a comparison of LDA and self-energy corrected band gaps with respective experimental values for a variety of materials. The underestimation within the LDA as well as the improvement by the GW approximation are evident.

Band gaps are experimentally measured by photoelectron spectroscopy. Figure 2 gives a schematic illustration. In direct photoelectron spectroscopy a photon with energy $\hbar\omega$ impinges on the sample and ejects an electron, whose kinetic energy $E_{\rm kin}$ is subsequently measured. The binding energy ϵ_i of this electron is given by the difference $\epsilon_i = E_{\rm kin} - \hbar\omega$. Actually, we already simplified the argumentation here, as the formulation "binding energy of an electron" suggests that the electrons are independent. In reality they are correlated through the Coulomb interaction, and the ejection of an electron is always a many-body process. In this general sense ϵ_i equals the difference $\epsilon_i = E_0^N - E_i^{N-1}$ between the total energy E_0^N of the N-particle ground state Ψ_0^N and the energy E_i^{N-1} of the $(N-1)$-particle state Ψ_i^{N-1} that remains after the emission. Inverse photoelectron spectroscopy is the complementary process: electrons are injected into the sample, and the energy of the emitted photon is measured. The number of electrons in the system thus increases from N to $N+1$, and we can identify $E_{\rm kin} - \hbar\omega$ with the energy difference $\epsilon_i = E_i^{N+1} - E_0^N$ of the many-electron systems.

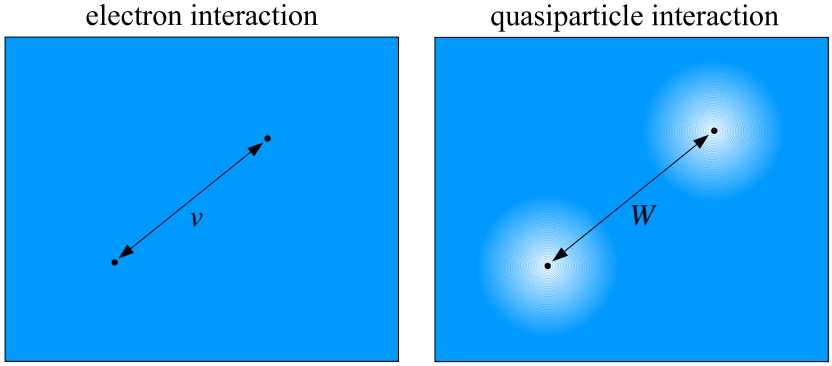


Fig. 3: *The electrons in a many-electron system are correlated by the strong Coulomb interaction v. The motion of one electron depends on the motion of all other electrons. A nearly-independent-particle picture can be recovered within the quasiparticle concept. Due to exchange and correlation effects a Coulomb hole forms around an electron and behaves together with it like a single entity, which is called quasiparticle. Quasiparticles interact via a weak screened interaction W instead of the strong Coulomb interaction.*

The fact that the independent-electron picture breaks down due to the strong Coulomb interaction questions single-electron concepts like band structure or Fermi surface. Still, in practice these work surprisingly well. In fact, we can at least retain a nearly-independent-particle picture if we consider quasiparticles instead of electrons (or holes). In the case of electron injection into a sample the repulsive Coulomb interaction creates a Coulomb hole around the additional electron (see Fig. 3). Analogously, if an electron leaves the system, its Coulomb hole also disappears. Relative to the ground-state N-electron system, the addition (removal) of an electron in indirect (direct) photoelectron spectroscopy hence creates (annihilates) an ensemble consisting of the bare electron and its oppositely charged Coulomb hole. This ensemble behaves in many ways like a single particle and is thus called "quasiparticle". Since the Coulomb hole reduces the total charge of the quasiparticle, the effective interaction between quasiparticles is screened and considerably weaker than the bare Coulomb interaction between electrons. In fact, the screened interaction is sufficiently small so that the quasiparticles can be regarded as approximately independent, which finally justifies the independent-particle approximation and explains the success of mean-field theories.

A theoretical description of processes involving the ejection or injection of electrons requires a framework that links the N-particle with the $(N \pm 1)$-particle systems. For this purpose we employ many-body perturbation theory. The central variable is the time-ordered Green function $G(\mathbf{r}t, \mathbf{r}'t')$. As we will see, it contains the excitation energies ϵ_i and even the excitation lifetimes. Besides, we can directly obtain the ground-state electron density, the expectation values of one-particle operators and the ground-state total energy from it. The Green function is hence capable of giving access to the same observables as the ground-state electron density. In contrast to the DFT expression $E\,[n]$, the functional $E\,[G]$ is even known exactly [2]. While the Green function contains much more information than the electron density, it is also a more complicated function and thus rarely applied to ground-state properties. In the present lecture we will, therefore, concentrate on the calculation of excited states.

Section 2 lays the theoretical foundations of the method. More complicated derivations are deferred to the appendix. The GW approximation is discussed in Section 3.1, and some aspects

of its numerical implementation are given in Section 3.2. As an illustration, Section 3.3 presents a number of selected applications. Section 4 contains the summary.

2 Theory

2.1 Green function

In this section we introduce the time-ordered Green function and examine its properties. We use the second-quantization formulation of quantum mechanics [2, 3]. For the present purpose it is sufficient to know that this formalism involves field operators $\hat{\psi}(\mathbf{r})$ and $\hat{\psi}^\dagger(\mathbf{r})$ that describe the annihilation and the creation of an electron at the position $\mathbf{r}$, respectively. We will not take spin dependence explicitly into account. If necessary, the spin quantum number can simply be added to the formulas by considering it to be part of the spatial coordinate $\mathbf{r}$.
The Green function $G^{\mathrm{e}}(\mathbf{r}t, \mathbf{r}'t')$ is defined such that $i\hbar G^{\mathrm{e}}(\mathbf{r}t, \mathbf{r}'t')$ is the probability amplitude for the propagation of an additional electron from $(\mathbf{r}', t')$ to $(\mathbf{r}, t)$ in a many-electron system with the Hamiltonian (41). This process brings the system from the N-electron ground state $\left|\Psi_0^N(t')\right\rangle$ to a final state $\hat{\psi}(\mathbf{r})U(t,t')\hat{\psi}^\dagger(\mathbf{r}')\left|\Psi_0^N(t')\right\rangle$. The final state is constructed by the successive action of the electron creation operator $\hat{\psi}^\dagger(\mathbf{r}')$, the evolution operator $\hat{U}(t,t') = \exp[-i\hat{H}(t-t')/\hbar]$, which takes the system from t' to a later time $t > t'$, and the electron annihilation operator $\hat{\psi}(\mathbf{r})$ on the N-electron ground state. As the probability amplitude is given by the overlap of the final state with $\left|\Psi_0^N(t)\right\rangle$, the Green function becomes

$$\begin{aligned} G^{\mathrm{e}}(\mathbf{r}t, \mathbf{r}'t') &= -\frac{i}{\hbar}\left\langle \Psi_0^N(t)\left|\hat{\psi}(\mathbf{r})\hat{U}(t,t')\hat{\psi}^\dagger(\mathbf{r}')\right|\Psi_0^N(t')\right\rangle \theta(t-t') \\ &= -\frac{i}{\hbar}\left\langle \Psi_0^N\left|\hat{\psi}(\mathbf{r}t)\hat{\psi}^\dagger(\mathbf{r}'t')\right|\Psi_0^N\right\rangle \theta(t-t')\,, \end{aligned} \tag{1}$$

where $\theta(t-t')$ is the Heaviside step function defined by

$$\theta(t-t') = \begin{cases} 1 & \text{if } t > t'\,, \\ 0 & \text{if } t < t'\,. \end{cases} \tag{2}$$

For the last equality in (1) we changed from the Schrödinger to the Heisenberg picture, where the expression is particularly simple. States and operators in the two pictures are related by

$$|\Psi_{\mathrm{H}}\rangle = \hat{U}(0,t)\,|\Psi_{\mathrm{S}}(t)\rangle \quad \text{and} \quad \hat{A}_{\mathrm{H}}(t) = \hat{U}(0,t)\hat{A}_{\mathrm{S}}\hat{U}(t,0)\,. \tag{3}$$

In the following we always omit the indices S and H. Similarly, we have the Green function

$$G^{\mathrm{h}}(\mathbf{r}'t', \mathbf{r}t) = -\frac{i}{\hbar}\left\langle \Psi_0^N\left|\hat{\psi}^\dagger(\mathbf{r}'t')\hat{\psi}(\mathbf{r}t)\right|\Psi_0^N\right\rangle \theta(t'-t) \tag{4}$$

for the propagation of an additional hole from $(\mathbf{r}, t)$ to $(\mathbf{r}', t')$. As a matter of convenience, we combine the two expressions in one time-ordered Green function

$$G(\mathbf{r}t, \mathbf{r}'t') = G^{\mathrm{e}}(\mathbf{r}t, \mathbf{r}'t') - G^{\mathrm{h}}(\mathbf{r}'t', \mathbf{r}t) = -\frac{i}{\hbar}\left\langle \Psi_0^N\left|\hat{T}\left[\hat{\psi}(\mathbf{r}t)\hat{\psi}^\dagger(\mathbf{r}'t')\right]\right|\Psi_0^N\right\rangle\,, \tag{5}$$

where we used the time-ordering operator $\hat{T}$, which rearranges a series of field operators in order of ascending time arguments from right to left with a factor (-1) for each pair permutation. Depending on the time order, Eq. (5) describes either electron ($t > t'$) or hole ($t < t'$) propagation. The electron density $n(\mathbf{r})$ can be expressed in terms of the Green function as

$$n(\mathbf{r}t) = \left\langle \Psi_0^N \left| \hat{\psi}^\dagger(\mathbf{r}t)\hat{\psi}(\mathbf{r}t) \right| \Psi_0^N \right\rangle = -i\hbar G(\mathbf{r}t, \mathbf{r}t+\eta)\,. \tag{6}$$

Here and in the following η is an infinitesimal positive number. It serves only to enforce the correct order of the field operators. Its unit should always be clear from the context; presently it is an infinitesimal time.

Let us consider the time-ordered Green function $G(\mathbf{r}, \mathbf{r}'; \tau)$ of a stationary system with $\tau = t - t'$. If we insert the closure relation $\sum_i \left|\Psi_i^{N\pm1}\right\rangle \left\langle\Psi_i^{N\pm1}\right| = 1$ between the two field operators in (5), where $\left\{\left|\Psi_i^{N\pm1}\right\rangle\right\}$ is the complete set of state vectors of the $(N \pm 1)$-particle system, transform to the Schrödinger picture and use the definitions

$$\psi_i^{N-1}(\mathbf{r}) = \left\langle \Psi_i^{N-1} \left| \hat{\psi}(\mathbf{r}) \right| \Psi_0^N \right\rangle \quad \text{and} \quad \psi_i^{N+1}(\mathbf{r}) = \left\langle \Psi_0^N \left| \hat{\psi}(\mathbf{r}) \right| \Psi_i^{N+1} \right\rangle \tag{7}$$

together with the excitation energies

$$\epsilon_i^{N-1} = E_0^N - E_i^{N-1} \quad \text{and} \quad \epsilon_i^{N+1} = E_i^{N+1} - E_0^N\,, \tag{8}$$

then we obtain

$$\begin{aligned} G(\mathbf{r}, \mathbf{r}'; \tau) \;=\; & -\frac{i}{\hbar} \sum_i \psi_i^{N+1}(\mathbf{r}) \psi_i^{N+1*}(\mathbf{r}') e^{-i\epsilon_i^{N+1}\tau/\hbar} \theta(\tau) \\ & + \frac{i}{\hbar} \sum_i \psi_i^{N-1}(\mathbf{r}) \psi_i^{N-1*}(\mathbf{r}') e^{-i\epsilon_i^{N-1}\tau/\hbar} \theta(-\tau)\,. \end{aligned} \tag{9}$$

The sums run over the ground state and all excited states of the $(N-1)$- and $(N+1)$-particle system, respectively. Expression (9) can be interpreted as follows: The state after the addition of an electron ($\tau > 0$) is represented by a linear combination of excited states

$$\hat{\psi}^\dagger(\mathbf{r}') \left|\Psi_0^N\right\rangle = \sum_i \psi_i^{N+1*}(\mathbf{r}') \left|\Psi_i^{N+1}\right\rangle \tag{10}$$

that subsequently evolve according to their respective phase factors $\exp(-i\epsilon_i^{N+1}\tau/\hbar)$. The resulting state is then probed at the point $\mathbf{r}$ by the projections $\psi_i^{N+1}(\mathbf{r})$. The case $\tau < 0$ (hole propagation) is analogous. Consequently, the Green function indeed contains the complete excitation spectrum of the $(N \pm 1)$-particle system. Fourier transformation of (9) to the frequency axis using the Fourier transform of the Heaviside step function

$$\theta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \theta(\tau)\; e^{i\omega\tau - \eta|\tau|} d\tau = \frac{i}{2\pi\,(\omega + i\eta)} \tag{11}$$

finally yields the Lehmann representation of the Green function

$$G(\mathbf{r}', \mathbf{r}; \omega) = \sum_i \frac{\psi_i^{N+1}(\mathbf{r}) \psi_i^{N+1*}(\mathbf{r})}{\hbar\omega - \epsilon_i^{N+1} + i\eta} + \sum_i \frac{\psi_i^{N-1}(\mathbf{r}) \psi_i^{N+1*}(\mathbf{r})}{\hbar\omega - \epsilon_i^{N-1} - i\eta}\,. \tag{12}$$

We observe that the Green function has poles at the true many-particle excitation energies $\epsilon_i^{N\pm1}$. These energies correspond to excitations of an $(N-1)$-particle and an $(N+1)$-particle system and hence to those processes measured in direct and inverse photoelectron spectroscopy. In the case of a noninteracting (or mean-field) system the $\psi_i^{N+1}(\mathbf{r})$ are simply the unoccupied and the $\psi_i^{N-1}(\mathbf{r})$ the occupied single-particle wave functions, the $\epsilon_i^{N\pm1}$ are the corresponding single-particle energies. In order not to overload the notation, we will drop the $(N \pm 1)$ superscripts from now on.

2.2 Spectral function

In connection with Eq. (9) we can define the spectral function $A(\mathbf{r}, \mathbf{r}'; \omega)$, i.e., the density of the excited (or quasiparticle) states that contribute to the electron or hole propagation. In a finite system this density is a series of delta functions at the excitation energies

$$A(\mathbf{r}, \mathbf{r}'; \omega) = \sum_i \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}')\delta(\hbar\omega - \epsilon_i) \tag{13}$$

weighted by the products of the corresponding projections (7). This allows us to rewrite (12) as an integral over frequencies

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \hbar \int_{-\infty}^{\infty} \frac{A(\mathbf{r}, \mathbf{r}'; \omega')}{\hbar\omega - \hbar\omega' + \mathrm{sgn}(\hbar\omega' - \mu)\, i\eta} d\omega' \tag{14}$$

with $\max\left(\epsilon_i^{N-1}\right) \leq \mu \leq \min\left(\epsilon_i^{N+1}\right)$. In an infinite system μ corresponds to the chemical potential. The inequality $\max\left(\epsilon_i^{N-1}\right) \leq \min\left(\epsilon_i^{N+1}\right)$ follows from the convexity of the total energy as a function of the particle number, i.e., $E_0^{N-1} - E_0^N \geq E_0^N - E_0^{N+1}$: we lose more energy when removing an electron than we gain by adding one. With the identity

$$\frac{1}{x \mp i\eta} = \mathcal{P}\left(\frac{1}{x}\right) \pm i\pi\delta(x) \tag{15}$$

in the limit $\eta \to 0^+$, where $\mathcal{P}(1/x)$ is the principal value of $1/x$, we find that

$$A(\mathbf{r}, \mathbf{r}'; \omega) = -\mathrm{sgn}(\hbar\omega - \mu)\, \frac{1}{\pi}\, \mathrm{Im}\, G(\mathbf{r}, \mathbf{r}'; \omega)\,. \tag{16}$$

The closure relation of the functions (7) yields another important property

$$\hbar \int_{-\infty}^{\infty} A(\mathbf{r}, \mathbf{r}'; \omega) d\omega = \sum_i \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}')\,. \tag{17}$$

When we change from a finite to an infinite system, the delta functions in $A(\mathbf{r}, \mathbf{r}'; \omega)$ merge and form a series of smooth peaks with finite line widths instead of sharp resonances (see Fig. 4). However, if the resulting spectral features are of Lorentzian form, i.e.,

$$A(\mathbf{r}, \mathbf{r}'; \omega) = \sum_i \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}') \frac{\Gamma_i}{\left(\omega - \tilde{\epsilon}_i/\hbar\right)^2 + \Gamma_i^2}\,, \tag{18}$$

where the $\tilde{\epsilon}_i$ are the peak positions and $|\Gamma_i|$ the corresponding peak widths, then we can perform the integration in (14) analytically and again obtain a discrete sum over i as in Eq. (12), provided

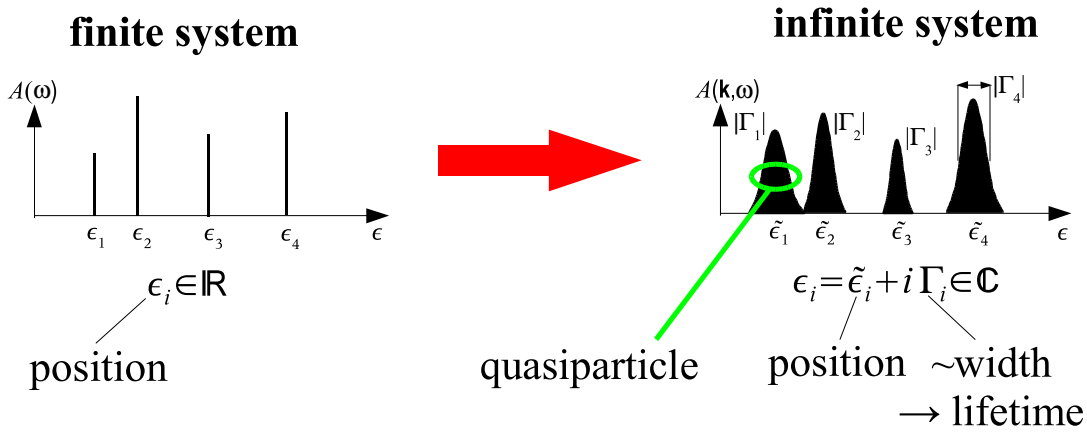


Fig. 4: *The excitation peaks of a finite system in the spectral function $A(\omega)$ merge into quasiparticle peaks of finite width in the case of an infinite system. This gives rise to finite excitation lifetimes determined by the inverse of the peak widths.*

that the energies are defined as complex numbers $\epsilon_i = \tilde{\epsilon}_i + i\Gamma_i$. Consequently, the form of the Fourier transform (9) remains unchanged, too. The imaginary component of ϵ_i leads to a damping term $\exp(-|\Gamma_i\tau|/\hbar)$, revealing that the excitation has a finite lifetime of $\hbar\,|\Gamma_i|^{-1}$. Physically, the de-excitation proceeds via Auger transitions that create electron-hole pairs on the way. The damping of the particle propagator G may seem surprising, as it suggests that the particle gradually disappears. However, one must keep in mind that we deal with an infinite system, i.e., $N \to \infty$, and an additional particle (electron or hole) can "dissipate" into the Fermi sea. In this sense, one often speaks of finite quasiparticle lifetimes and calls $\psi_i(\mathbf{r})$ and ϵ_i the quasiparticle wave functions and energies, respectively. The quasiparticle equation (22) introduced in the next section holds for infinite systems if one uses an analytic continuation of the self-energy into the complex frequency plane.

2.3 Dyson equation

Appendix A shows that the time-ordered Green function $G(\mathbf{r},\mathbf{r}';\omega)$ of the interacting system obeys an integral equation, the Dyson equation

$$G(\mathbf{r},\mathbf{r}';\omega) = G_0(\mathbf{r},\mathbf{r}';\omega) + \iint G_0(\mathbf{r},\mathbf{r}'';\omega)\Sigma(\mathbf{r}'',\mathbf{r}''';\omega)G(\mathbf{r}''',\mathbf{r}';\omega)d^3r''d^3r''' \,, \tag{19}$$

where $G_0(\mathbf{r},\mathbf{r}';\omega)$ is the Green function of a mean-field system defined by

$$\hat{h}_0\varphi_i^0(\mathbf{r}) = \epsilon_i^0\varphi_i^0(\mathbf{r}) \tag{20}$$

with the single-particle Hamiltonian

$$\hat{h}_0(\mathbf{r}) = -\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + \frac{e^2}{4\pi\varepsilon_0}\int\frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3r' \,. \tag{21}$$

The quantities $V_{\text{ext}}(\mathbf{r})$, m, e and ε_0 are defined as in Eq. (41). The Green function $G_0(\mathbf{r},\mathbf{r}';\omega)$ is obtained from Eq. (12) with the wave functions $\varphi_i^0(\mathbf{r})$ and energies ϵ_i^0. The nonlocal and

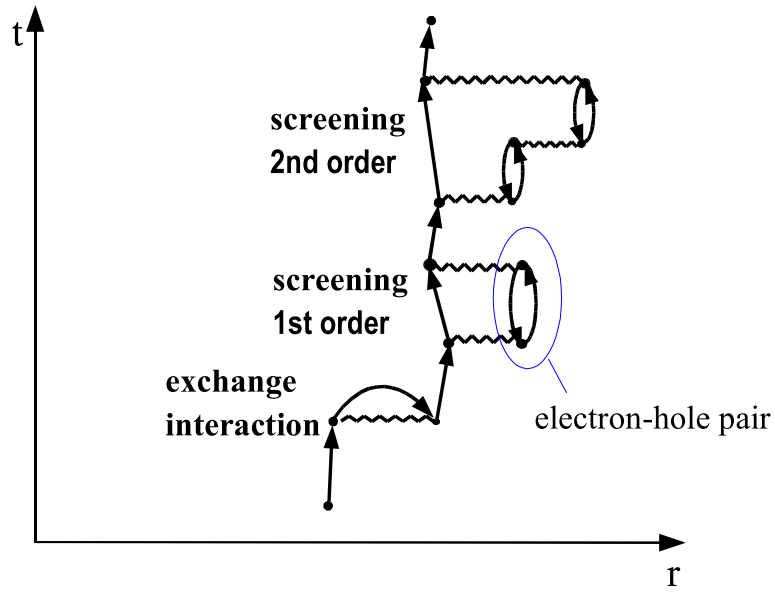


Fig. 5: *Illustration of a series of scattering processes using Feynman diagrams. All zigzag lines representing the instantaneous Coulomb interaction must be drawn horizontally. Arrows going forward in time represent electron and those going backward in time hole propagators. The self-energy is the sum of all possible single scattering processes.*

frequency-dependent function $\Sigma(\mathbf{r}, \mathbf{r}', \omega)$ is the non-Hermitian self-energy operator, which contains all many-body exchange and correlation effects beyond the electrostatic Hartree potential. This can be more easily seen in a reformulation of the Dyson equation. By inserting the Lehmann representation (12) into Eq. (19), we find that the wave functions $\psi_i(\mathbf{r})$ and energies ϵ_i obey the quasiparticle equation

$$\hat{h}_0(\mathbf{r})\psi_i(\mathbf{r}) + \int \Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_i/\hbar)\psi_i(\mathbf{r}')d^3r' = \epsilon_i\psi_i(\mathbf{r}) \tag{22}$$

(see appendix B), which is nonlinear in ϵ_i. Although it looks very similar to the one-particle equations of mean-field approaches like Hartree, Hartree-Fock or DFT, it does *not* constitute a mean-field formulation, since the self-energy takes all dynamic many-electron processes into account. Consequently, the functions $\psi_i(\mathbf{r})$ and energies ϵ_i must not be understood as single-particle quantities. In fact, they are defined in Eqs. (7) and (8) as properties of the many-electron system. From the nonlinearity of the quasiparticle equation it follows that the wave functions $\psi_i(\mathbf{r})$ are not orthonormal, in contrast to single-particle wave functions. However, they do fulfill the closure relation (17).
The Dyson equation (19) can be rewritten in the form of a geometric series by subsequently replacing G on the right-hand side by $G_0 + G_0\Sigma G$, which leads to, symbolically written,

$$G = G_0 + G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0\Sigma G_0 + \dots . \tag{23}$$

This is a typical equation of scattering theory, where the different terms of the geometric series describe single, double, triple, etc., scattering processes, and Σ is the scattering potential. Such a succession of scattering processes can be illustrated by Feynman diagrams, where G_0 is drawn as a straight arrow and the Coulomb interaction as a zigzag line. According to (23), a

diagrammatic representation of a multiple scattering process should involve a series of arrows (G_0) divided by single scattering processes (Σ). In the example of Fig. 5 these are the exchange interaction, the creation of an electron-hole pair (the "bubble" diagram) and finally the creation of a pair that itself creates another pair. In order to obtain the complete Green function, we have to sum all multiple scattering processes, of which the one shown in Fig. 5 is merely one example. The self-energy is given by the sum of all single scattering processes. The interpretation in terms of scattering processes allows to construct approximations for Σ by the summation of diagrams considered essential for the physical behavior of a given electron system. In general, however, such approximations are rarely convergent, and too many processes turn out to be quantitatively important. Therefore, we apply a systematic algebraic method instead.

3 Implementation and Applications

3.1 *GW* approximation

In practice we must use an approximation for the self-energy, such as the GW approximation, which contains the electron exchange and a large part of the electron correlation. It is formally derived in Appendix A and has a very simple mathematical form in the time domain

$$\Sigma^{GW}(\mathbf{r},\mathbf{r}';\tau) = i\hbar G_0(\mathbf{r},\mathbf{r}';\tau)W(\mathbf{r},\mathbf{r}';\tau+\eta)\,. \tag{24}$$

In order to calculate the self-energy contribution to the quasiparticle energies, we need the Fourier transform on the frequency axis

$$\Sigma^{GW}(\mathbf{r},\mathbf{r}';\omega) = \frac{i\hbar}{2\pi}\int_{-\infty}^{\infty} G_0(\mathbf{r},\mathbf{r}';\omega+\omega')W(\mathbf{r},\mathbf{r}';\omega')e^{i\omega'\eta}d\omega'\,. \tag{25}$$

The first function on the right-hand side is the Green function of the noninteracting system defined by (20) and the second function the dynamically screened interaction $W(\mathbf{r},\mathbf{r}';\omega)$, which is related to the bare Coulomb potential $v(\mathbf{r},\mathbf{r}') = e^2/\left(4\pi\varepsilon_0\left|\mathbf{r}-\mathbf{r}'\right|\right)$ through the inverse of the dielectric function

$$W(\mathbf{r},\mathbf{r}';\omega) = \int \varepsilon^{-1}(\mathbf{r},\mathbf{r}'';\omega)v(\mathbf{r}'',\mathbf{r}')d^3r'' = v(\mathbf{r},\mathbf{r}') + \int n_{\mathrm{ind}}(\mathbf{r},\mathbf{r}'';\omega)v(\mathbf{r}'',\mathbf{r}')d^3r''\,. \tag{26}$$

The screened interaction $W(\mathbf{r},\mathbf{r}';\omega)$ is the effective potential at $\mathbf{r}'$ induced by a quasiparticle at $\mathbf{r}$: the Coulomb potential of the electron repels other electrons in its neighborhood and thus gives rise to the formation of an exchange and correlation hole, whose effective positive charge $n_{\mathrm{ind}}(\mathbf{r},\mathbf{r}'';\omega)$ screens the bare Coulomb potential $v(\mathbf{r},\mathbf{r}')$ (see Fig. 6). Analogously, an effective negative charge screens the Coulomb potential of a hole. The screened interaction is considerably weaker than the bare Coulomb interaction. The GW approximation uses the random-phase approximation (RPA)

$$\varepsilon(\mathbf{r},\mathbf{r}';\omega) = \delta(\mathbf{r}-\mathbf{r}') - \int v(\mathbf{r},\mathbf{r}'')P(\mathbf{r}'',\mathbf{r}';\omega)d^3r''\,, \tag{27}$$

$$P(\mathbf{r},\mathbf{r}';\tau) = -i\hbar G_0(\mathbf{r},\mathbf{r}';\tau)G_0(\mathbf{r}',\mathbf{r};-\tau)\,. \tag{28}$$

It corresponds to a subset of scattering processes in the many-electron system. Some of the respective diagrams are just the ones shown in Fig. 5. Using expression (12) for the Green

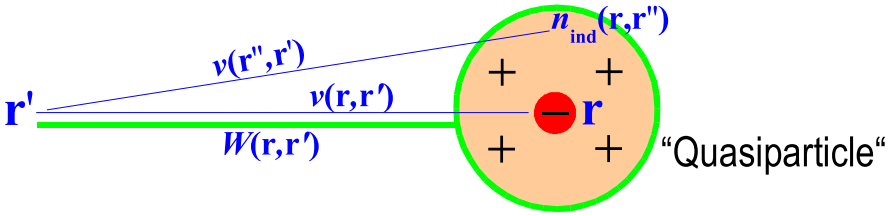


Fig. 6: *The formation of the Coulomb hole around an electron at* $\mathbf{r}$ *screens its Coulomb potential* $v(\mathbf{r},\mathbf{r}')$*. This leads to the definition of the screened interaction* $W(\mathbf{r},\mathbf{r}')$ *that takes into account the combined potentials of the bare electron and its screening cloud* n_{ind}*. The ensemble consisting of the electron and its polarization cloud is called "quasiparticle".*

function G_0 of the noninteracting system we observe that the Fourier transform of the polarization function (28) is given by

$$P(\mathbf{r},\mathbf{r}';\omega) = \sum_{i}^{\text{occ.}} \sum_{j}^{\text{unocc.}} \varphi_i^0(\mathbf{r})\,\varphi_j^{0*}(\mathbf{r}) \left(\frac{1}{\hbar\omega + \epsilon_i^0 - \epsilon_j^0 + i\eta} - \frac{1}{\hbar\omega - \epsilon_i^0 + \epsilon_j^0 - i\eta} \right) \varphi_i^{0*}(\mathbf{r}')\,\varphi_j^0(\mathbf{r}') \tag{29}$$

in terms of the wave functions $\varphi_i^0(\mathbf{r})$ and energies ϵ_i^0.
The well-known Hartree-Fock equations can be recovered from Eq. (22) if we use the energy-independent self-energy

$$\Sigma^{\mathrm{HF}}(\mathbf{r},\mathbf{r}') = i\hbar G_0(\mathbf{r},\mathbf{r}';-\eta)v(\mathbf{r},\mathbf{r}') \tag{30}$$

(given in the time domain) instead. By comparison with Eq. (24), we see that the GW approximation constitutes an expansion of the self-energy up to first order in the screened interaction as opposed to the bare Coulomb interaction in (30). This approximates the exact self-energy considerably better, because W is much smaller than v. Due to the similarity of the two self-energy expressions, the GW approximation can formally be regarded as a Hartree-Fock approach with a dynamically screened interaction W instead of the static Coulomb interaction v.

3.2 Numerical implementation

For band-structure calculations it is more efficient to obtain the ϵ_i directly from the quasiparticle equation (22) instead of solving the Dyson integral equation (19) and searching for the poles of the Green function. Furthermore, it is then possible to exploit the formal similarity to the Kohn-Sham equation

$$\hat{h}_0\varphi_i^{\mathrm{KS}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})\varphi_i^{\mathrm{KS}}(\mathbf{r}) = \epsilon_i^{\mathrm{KS}}\varphi_i^{\mathrm{KS}}(\mathbf{r})\,, \tag{31}$$

where $V_{\mathrm{xc}}(\mathbf{r})$ is the local exchange-correlation potential. In many cases the Kohn-Sham eigenvalues ϵ_i^{KS} already provide a reasonable estimate of the band structure and are in qualitative agreement with experiment. For systems where the quasiparticle wave functions are known, one also finds $\varphi_i^{\mathrm{KS}}(\mathbf{r}) \approx \psi_i(\mathbf{r})$ [9]. This observation indicates that the self-energy correction $\Sigma(\mathbf{r},\mathbf{r}';\epsilon_i/\hbar) - V_{\mathrm{xc}}(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')$ is small and justifies the use of first-order perturbation theory to obtain approximate energies

$$\epsilon_i \approx \epsilon_i^{\mathrm{KS}} + \left\langle \varphi_i^{\mathrm{KS}} \middle| \Sigma(\epsilon_i/\hbar) - V_{\mathrm{xc}} \middle| \varphi_i^{\mathrm{KS}} \right\rangle\,. \tag{32}$$

A solution of this nonlinear equation still requires the knowledge of the frequency dependence of the self-energy, which is not known in general. Therefore, we use the linear expansion

$$\Sigma(\mathbf{r},\mathbf{r}';\epsilon_i/\hbar) \approx \Sigma(\mathbf{r},\mathbf{r}';\epsilon_i^{\mathrm{KS}}/\hbar) + \frac{\epsilon_i - \epsilon_i^{\mathrm{KS}}}{\hbar} \frac{\partial\Sigma(\mathbf{r},\mathbf{r}';\epsilon_i^{\mathrm{KS}}/\hbar)}{\partial\omega}\,, \tag{33}$$

which leads to

$$\epsilon_i \approx \epsilon_i^{\mathrm{KS}} + Z_i \left\langle \varphi_i^{\mathrm{KS}} \left| \Sigma\left(\epsilon_i^{\mathrm{KS}}/\hbar\right) - V_{\mathrm{xc}} \right| \varphi_i^{\mathrm{KS}} \right\rangle . \tag{34}$$

The quasiparticle renormalization factor is given by

$$Z_i = \left(1 - \left\langle \varphi_i^{\mathrm{KS}} \left| \frac{1}{\hbar} \frac{\partial\Sigma\left(\epsilon_i^{\mathrm{KS}}/\hbar\right)}{\partial\omega} \right| \varphi_i^{\mathrm{KS}} \right\rangle \right)^{-1} \tag{35}$$

and equals the quasiparticle weight

$$Z_i = \int |\psi_i(\mathbf{r})|^2 \, d^3r < 1. \tag{36}$$

With the decomposition of W into the bare Coulomb interaction v and the remainder $W - v$, the GW self-energy (24) splits into exchange and correlation parts, symbolically written as

$$\Sigma^{GW} = i\hbar G_0^{\mathrm{KS}} W = i\hbar G_0^{\mathrm{KS}} v + i\hbar G_0^{\mathrm{KS}} (W - v) = \Sigma_{\mathrm{x}}^{GW} + \Sigma_{\mathrm{c}}^{GW} . \tag{37}$$

Instead of G_0 we use the Kohn-Sham Green function G_0^{KS}. After inserting this decomposition into Eq. (25), we must evaluate the convolutions

$$\Sigma_{\mathrm{x}}^{GW}(\mathbf{r},\mathbf{r}';\omega) = \frac{i\hbar}{2\pi} \int_{-\infty}^{\infty} G_0^{\mathrm{KS}}(\mathbf{r},\mathbf{r}';\omega+\omega') v(\mathbf{r},\mathbf{r}') e^{i\omega'\eta} d\omega' \,, \tag{38a}$$

$$\Sigma_{\mathrm{c}}^{GW}(\mathbf{r},\mathbf{r}';\omega) = \frac{i\hbar}{2\pi} \int_{-\infty}^{\infty} G_0^{\mathrm{KS}}(\mathbf{r},\mathbf{r}';\omega+\omega') \left[W(\mathbf{r},\mathbf{r}';\omega') - v(\mathbf{r},\mathbf{r}') \right] d\omega' \,. \tag{38b}$$

The integral (38a) can be evaluated analytically and leads to the well-known expression for the Hartree-Fock exchange term

$$\left\langle \varphi_i^{\mathrm{KS}} \left| \Sigma_{\mathrm{x}}^{GW} \right| \varphi_i^{\mathrm{KS}} \right\rangle = -\frac{e^2}{4\pi\varepsilon_0} \sum_j^{\mathrm{occ.}} \int \frac{\varphi_i^{\mathrm{KS}*}(\mathbf{r})\, \varphi_j^{\mathrm{KS}}(\mathbf{r})\, \varphi_j^{\mathrm{KS}*}(\mathbf{r}')\, \varphi_i^{\mathrm{KS}}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d^3r\, d^3r' \,. \tag{39}$$

In general, the second convolution (38b) must be computed numerically. For this purpose the integration contour is usually deformed to the complex plane, where the analytical continuations of G_0 and W are smoother.

Let the Kohn-Sham wave functions be represented in a basis $\{\zeta_\alpha(\mathbf{r})\}$. According to (29) we can then write the polarization function and all related quantities in terms of products $\chi_\mu(\mathbf{r}) = \zeta_\alpha^*(\mathbf{r})\zeta_\beta(\mathbf{r})$ with the composite index $\mu = (\alpha,\beta)$ as

$$P(\mathbf{r},\mathbf{r}';\omega) = \sum_{\mu,\nu} P_{\mu\nu}(\omega) \chi_\mu^*(\mathbf{r}) \chi_\nu(\mathbf{r}') \,. \tag{40}$$

The Eqs. (24) to (28) are solved by matrix operations:

1. A self-consistent DFT loop produces the Kohn-Sham wave functions $\varphi_i^{\mathrm{KS}}(\mathbf{r})$ and energies ϵ_i^{KS}. At this point we can already evaluate the exchange term (39).

2. The polarization matrix $P_{\mu\nu}(\omega)$ is calculated according to (29).
3. The dielectric matrix is obtained from $\varepsilon_{\mu\nu}(\omega) = \delta_{\mu\nu} - \sum_\gamma v_{\mu\gamma} P_{\gamma\nu}(\omega)$ and inverted.
4. Next the screened interaction $W_{\mu\nu}(\omega) = \sum_\gamma \varepsilon^{-1}_{\mu\gamma}(\omega) v_{\gamma\nu}$ is calculated from a matrix multiplication of the inverse dielectric function with the Coulomb matrix.
5. The correlation term $\left\langle \varphi_i^{\mathrm{KS}} \left| \Sigma_{\mathrm{c}}^{GW} \right| \varphi_i^{\mathrm{KS}} \right\rangle$ is evaluated according to (38b) with a numerical contour integration on the complex frequency plane.
6. Finally, approximate quasiparticle energies are obtained from (34) and (35).

The computation of the dielectric function, its inversion and the convolution (38b) are very time-consuming. Therefore, some (especially older) codes approximate the inverse dielectric function by a so-called plasmon-pole model [10, 11]. These models replace the imaginary component of $\varepsilon^{-1}(\omega)$, which has a peaked structure, by a sum of delta functions at the corresponding frequencies. This simplification reduces the third step to a single matrix inversion of the static dielectric function at $\omega = 0$ and makes an analytic evaluation of the frequency integral (38b) possible.

3.3 Examples

Although Hedin's seminal article [4] was already published in 1965, it was not before the middle of the 1980s that the first ab initio calculations for real materials were reported in the literature.

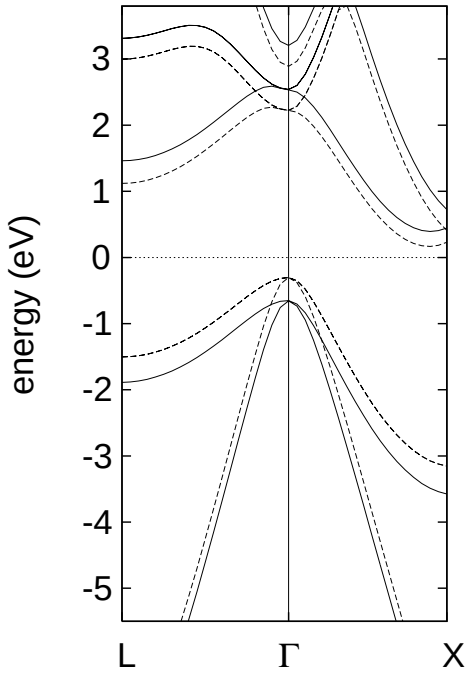


Fig. 7: *LDA band structure (dashed lines) of silicon with GW self-energy corrected valence and conduction bands (solid lines). The GW approximation shifts the corresponding bands up and down, respectively, but leaves the dispersion essentially unaffected.*

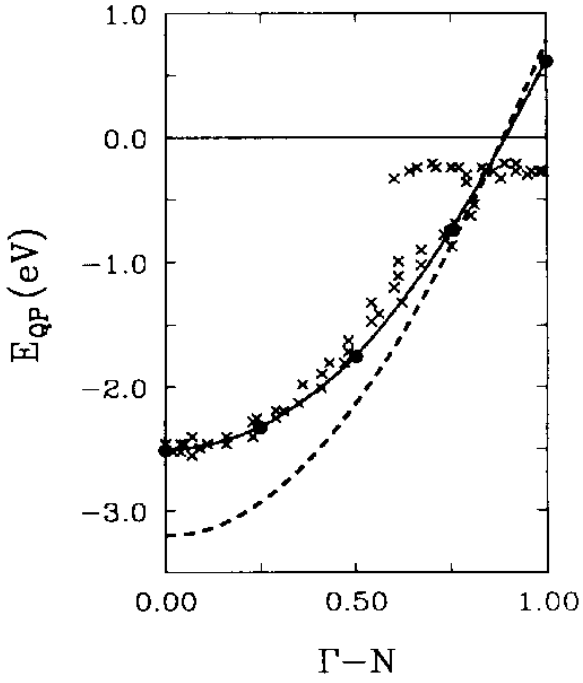


Fig. 8: *Comparison of LDA (dashed), quasiparticle (solid line) and experimental (crosses) bands for Na. Taken from Ref. [16].*

In spite of several approximations in the numerical treatment, which were necessary because of the lack of computer power back then, initial results were already very promising. Hybertsen and Louie [12] as well as Godby *et al.* [13] showed that the calculated band gap of Si fell within a margin of about 0.1 eV from the experimental value. Shortly afterwards the same authors reported band gaps for several other semiconducting materials that turned out to be equally accurate [14, 15]. After these pioneering studies the GW approximation was applied to a variety of semiconductors with great success (see, e.g., Fig. 1). The principal effect of the GW self-energy correction on the band structure of a semiconductor is to rigidly shift the valence bands up and the conduction bands down, thus opening the band gap. Figure 7 shows this effect for Si as an example.

Not only the band gaps of semiconductors and insulators are improved by the GW self-energy correction, but the correlation-induced band narrowing of metals is also correctly described. The band narrowing reflects the higher effective mass of quasiparticles (the polarization cloud adds to the electron mass) compared to bare electrons. For this reason, the self-energy is sometimes also referred to as "mass operator". Figure 8 shows the energy dispersion of Na as an example [16]. The band narrowing brought about by the GW self-energy correction leads to nearly perfect agreement with experiment.

The calculated excitation or quasiparticle lifetimes can be directly compared with two-photon photoemission spectroscopy. This experimental method allows to measure dynamical de-excitation processes in electronic systems. After a first photon has excited the electron system (creating a "hot" electron), a second photon probes the quasiparticle density of states like in ordinary direct photoelectron spectroscopy. The time delay between the two photons can be tuned such that the system can be observed in different stages of the electronic de-excitation process. From a series of measurements one can thus deduce the lifetime, which depends on

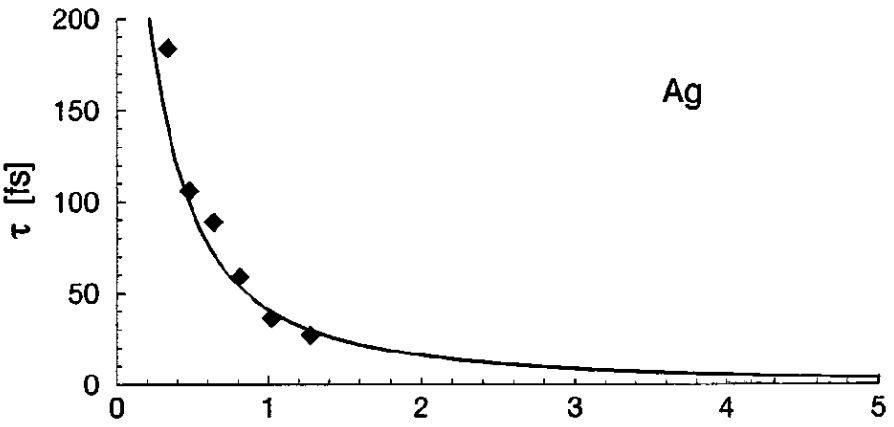


Fig. 9: *De-excitation dynamics measured in time-resolved two-photon photoemission spectroscopy (diamonds) and calculated with the GW approximation (solid line). Taken from Ref. [17].*

the excitation energy, i.e., the energy of the first photon. In the example of Ag in Fig. 9 the theoretical curve $\tau = \hbar\,|\Gamma|^{-1}$ obtained from the imaginary parts of the quasiparticle energies (see Sec. 2.2) closely follows the experimental data points [17].

4 Summary

In this lecture we presented the GW approximation for the electronic self-energy, which allows to calculate excited-state properties like excitation energies and lifetimes. The self-energy describes scattering processes between electrons and, in principle, contains all exchange and correlation effects beyond the electrostatic Hartree potential. The GW approximation includes a subset of these scattering processes. Apart from exchange it describes the creation of electron-hole pairs within the random-phase approximation (RPA) that leads to the formation of polarization clouds around the bare particles. The ensemble of an electron or a hole together with its polarization cloud behaves essentially like a single entity and is called a quasiparticle. The quasiparticles interact via a screened potential that is considerably weaker than the bare Coulomb interaction. This makes a perturbative treatment possible. In this respect, the GW approximation constitutes an expansion of the self-energy up to linear order in the screened interaction. It works well in a large class of systems where the polarization effects covered by the RPA play the dominant role in electron correlation, such as simple metals and semiconductors. The GW approximation is by nature a perturbative approach. Actual GW calculations are usually based on the self-consistent Kohn-Sham wave functions and energies as a starting point. This method has its limitations in materials where DFT already gives unphysical results. It breaks down for systems with very strong electronic correlation, which is insufficiently described by the available exchange-correlation functionals. The large error in the band gap of NiO in Fig. 1 is an example. In reality, NiO is a strongly correlated Mott-Hubbard insulator, whereas it comes out as a semiconductor with a very small band gap (nearly a semi-metal) in DFT calculations.
The GW method is designed for the analysis of excited states of the $(N \pm 1)$-electron systems. The treatment of optical absorption processes, where the particle number does not change due

to the promotion of valence electrons into unoccupied conduction states rather than emission, requires the simultaneous description of two particles, an electron and a hole, i.e., an exciton. Consequently, one must describe such a process with a two-particle Green function. In this case many-body perturbation theory leads to the so-called Bethe-Salpeter equation. Absorption spectra obtained from this equation are indeed very accurate [18]. An alternative is time-dependent density-functional theory [19], which also gives access to the excited states of an N-electron system.

Appendices

A Hedin Equations

With the field-operators introduced in Section 2.1 we can rewrite the many-particle Hamiltonian

$$\hat{H} = \sum_i \left[-\frac{\hbar^2}{2m} \nabla_i^2 + V_{\mathrm{ext}}(\mathbf{r}_i) \right] + \frac{1}{2} \sum_{ij} v(\mathbf{r}_i, \mathbf{r}_j) \,, \tag{41}$$

where $V_{\mathrm{ext}}(\mathbf{r})$ is the potential created by the atomic nuclei, $v(\mathbf{r}, \mathbf{r}') = e^2 / (4\pi\varepsilon_0 |\mathbf{r} - \mathbf{r}'|)$ the Coulomb interaction, m the electron mass, e the electron charge and ε_0 the vacuum dielectric constant, as

$$\hat{H} = \int \hat{\psi}^\dagger(\mathbf{r}) \hat{h}(\mathbf{r}) \hat{\psi}(\mathbf{r}) d^3r + \frac{1}{2} \iint \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}^\dagger(\mathbf{r}') v(\mathbf{r}, \mathbf{r}') \hat{\psi}(\mathbf{r}') \hat{\psi}(\mathbf{r}) d^3r \, d^3r' \tag{42}$$

with the one-particle operator

$$\hat{h}(\mathbf{r}) = -\frac{\hbar^2}{2m} \nabla^2 + V_{\mathrm{ext}}(\mathbf{r}) \,. \tag{43}$$

The expression (42) is just a mathematical reformulation of (41) and should not be mistaken for the energy expectation value in Hartree theory, although it looks similar.
From the equation of motion for the annihilation operator

$$i\hbar \frac{\partial}{\partial t} \hat{\psi}(\mathbf{r}, t) = \left[\hat{\psi}(\mathbf{r}, t), \hat{H} \right]_- = \hat{h}(\mathbf{r}) \hat{\psi}(\mathbf{r}, t) + \int v(\mathbf{r}, \mathbf{r}') \hat{\psi}^\dagger(\mathbf{r}', t) \hat{\psi}(\mathbf{r}', t) \hat{\psi}(\mathbf{r}, t) d^3r' \,, \tag{44}$$

which describes the time evolution of a Heisenberg operator in the same way as the Schrödinger equation describes that of a wave function, we can directly deduce the equation of motion for the Green function

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} G(\mathbf{r}t, \mathbf{r}'t') &= \delta(\mathbf{r} - \mathbf{r}') \delta(t - t') + \hat{h}(\mathbf{r}) G(\mathbf{r}t, \mathbf{r}'t') \\ &\quad - \frac{i}{\hbar} \int v(\mathbf{r}, \mathbf{r}'') \left\langle \Psi_0^N \left| \hat{T} \left[\hat{\psi}^\dagger(\mathbf{r}'', t) \hat{\psi}(\mathbf{r}'', t) \hat{\psi}(\mathbf{r}, t) \hat{\psi}^\dagger(\mathbf{r}', t') \right] \right| \Psi_0^N \right\rangle d^3r'' \,. \end{aligned} \tag{45}$$

This is not a closed equation, because it involves the two-particle Green function

$$G(1234) = -\frac{1}{\hbar^2} \left\langle \Psi_0^N \left| \hat{T} \left[\hat{\psi}(1) \hat{\psi}(2) \hat{\psi}^\dagger(4) \hat{\psi}^\dagger(3) \right] \right| \Psi_0^N \right\rangle \,. \tag{46}$$

Here and in the following we denote the set of space-time coordinates $(\mathbf{r}_1, t_1)$ with a natural number 1, etc., and further define

$$\begin{aligned} \delta(12) &= \delta(\mathbf{r}_1 - \mathbf{r}_2)\delta(t_1 - t_2)\,, & (47)\\ v(12) &= v(\mathbf{r}_1, \mathbf{r}_2)\delta(t_1 - t_2)\,, & (48)\\ \int d1 &= \int d^3 r_1 \int_{-\infty}^{\infty} dt_1\,, & (49)\\ 1^+ &= (\mathbf{r}_1, t_1 + \eta)\,, & (50) \end{aligned}$$

where η is an infinitesimal positive time. With the two-particle Green function (46) we can rewrite (45) as

$$i\hbar \frac{\partial}{\partial t_1} G(12) = \delta(12) + \hat{h}(1)G(12) - i\hbar \int v(1^+3)G(1323^+)d3\,. \tag{51}$$

The additional infinitesimals in 1^+ and 3^+ make sure that the time order is the same as in (45). In order to employ the functional-derivative method, we introduce an external potential $U(1)$ that is again set to zero at the end. Of course, all quantities from now on depend on $U(1)$, while the equations remain invariant provided that we replace $\hat{h}(1) \to \hat{h}(1) + U(1)$. We can use functional differentiation to define a number of useful quantities. The reaction of the density to changes in the external potential is governed by the linear-response function

$$R(12) = \left.\frac{\delta n(1)}{\delta U(2)}\right|_{U=0}\,. \tag{52}$$

The test potential and the Coulomb potential created by the induced charge can be combined into an effective potential

$$U_{\text{eff}}(1) = U(1) + \iint v(13)R(32)U(2)d2\,d3\,, \tag{53}$$

which is related to $U(1)$ via the inverse dielectric function

$$\varepsilon^{-1}(12) = \left.\frac{\delta U_{\text{eff}}(1)}{\delta U(2)}\right|_{U=0} = \delta(12) + \int v(13)R(32)d3\,. \tag{54}$$

With the definition of the polarization function

$$P(12) = \left.\frac{\partial n(1)}{\partial U_{\text{eff}}(2)}\right|_{U=0} \tag{55}$$

and the chain rule for functional derivatives one obtains the geometric series

$$\varepsilon^{-1}(12) = \delta(12) + \int v(13)P(32)d3 + \iiint v(13)P(34)v(45)P(52)d3\,d4\,d5 + \ldots\,, \tag{56}$$

which can easily be inverted to yield

$$\varepsilon(12) = \delta(12) - \int v(13)P(32)d3\,. \tag{57}$$

If we take the Coulomb potential of an electron at 2 as the test potential, we get the screened potential

$$W(12) = \int \varepsilon^{-1}(13)v(32)d3 = v(12) + \iint v(13)P(34)W(42)d3\,d4 \tag{58}$$

as the effective potential at position 1.
After this interlude we can go on with the derivation. For the functional derivative of the Green function with respect to the test potential we find [4, 20]

$$\left.\frac{\delta G(12)}{\delta U(3)}\right|_{U=0} = G(12)G(33^+) - G(1323^+)\,. \tag{59}$$

This allows us to eliminate the two-particle Green function, and the integral in (51) becomes

$$\begin{aligned} -i\hbar \int v(1^+3)G(1323^+)d3 &= \underbrace{-i\hbar\left(\int v(13)G(33^+)d3\right)}_{V^{\mathrm{H}}(1)} G(12) + i\hbar \int v(1^+3)\frac{\delta G(12)}{\delta U(3)}d3 \\ &= V^{\mathrm{H}}(1)G(12) + \int \Sigma(13)G(32)d3\,, \end{aligned} \tag{60}$$

where $V^{\mathrm{H}}(1)$ is the Hartree potential [cf. Eq. (21)] and

$$\begin{aligned} \Sigma(12) &= i\hbar \iint v(1^+3)\frac{\delta G(14)}{\delta U(3)}G^{-1}(42)d3\,d4 \\ &= -i\hbar \iint v(1^+3)G(14)\frac{\delta G^{-1}(42)}{\delta U(3)}d3\,d4 \\ &= i\hbar \iint W(1^+3)G(14)\Gamma(42;3)d3\,d4 \end{aligned} \tag{61}$$

the self-energy. For the second identity we used partial integration and for the third the chain rule for functional derivatives, the definition of the screened interaction (58) and the vertex function

$$\Gamma(12;3) = -\left.\frac{\delta G^{-1}(12)}{\delta U_{\mathrm{eff}}(3)}\right|_{U=0}. \tag{62}$$

With the self-energy (61) the equation of motion for the Green function (51) now becomes

$$\left[i\hbar\frac{\partial}{\partial t_1} - \hat{h}_0(1)\right] G(12) - \int \Sigma(13)G(32)d3 = \delta(12)\,, \tag{63}$$

where we incorporated the Hartree potential into the one-particle operator

$$\hat{h}_0(1) = \hat{h}(1) + V^{\mathrm{H}}(1)\,. \tag{64}$$

The delta function on the right-hand side of Eq. (63) demonstrates that $G(12)$ is indeed a Green function in the mathematical sense. In a noninteracting system the self-energy vanishes, and (63) becomes

$$\left[i\hbar\frac{\partial}{\partial t_1} - \hat{h}_0(1)\right] G_0(12) = \delta(12)\,. \tag{65}$$

Multiplication of Eq. (63) with G_0 and Eq. (65) with G from the left followed by integration yields the Dyson equation

$$G(12) = G_0(12) + \iint G_0(13)\Sigma(34)G(42)d3\,d4\,. \tag{66}$$

Finally, Eqs. (65) and (66) allow us to rewrite the vertex function (62) as

$$\Gamma(12;3) = \delta(12)\delta(13) + \frac{\delta\Sigma(12)}{\delta U_{\text{eff}}(3)}\,, \tag{67}$$

and with the identity

$$\begin{aligned}\frac{\delta G(12)}{\delta U_{\text{eff}}(3)} &= \frac{\delta}{\delta U_{\text{eff}}(3)} \iint G(14)G^{-1}(45)G(52)d4\,d5 \\ &= 2\frac{\delta G(12)}{\delta U_{\text{eff}}(3)} + \iint G(14)\frac{\delta G^{-1}(45)}{\delta U_{\text{eff}}(3)}G(52)d4\,d5\end{aligned} \tag{68}$$

we obtain

$$\Gamma(12;3) = \delta(12)\delta(13) - \iiiint \frac{\delta\Sigma(12)}{\delta G(45)}G(56)\Gamma(67;3)G(74)d4\,d5\,d6\,d7 \tag{69}$$

and analogously

$$P(12) = -i\hbar\frac{\delta G(11^+)}{\delta U_{\text{eff}}(2)} = -i\hbar \iint G(13)\Gamma(34;2)G(41)d3\,d4\,. \tag{70}$$

The Eqs. (58), (66), (61), (69) and (70) constitute Hedin's set of integro-differential equations, whose self-consistent solution, in principle, solves the many-electron problem exactly. Unfortunately, they are not just numerical relations but contain a functional derivative in (69). Therefore, the Hedin equations cannot be solved self-consistently by computer codes, but they may be iterated analytically in order to derive useful approximations. In practice we can only perform one iteration. We start with the Green function G_0 of the noninteracting system, which corresponds to the single-particle Hamiltonian (64). As the corresponding self-energy vanishes in this case, the set of equations simplifies to

$$\Gamma(12;3) = \delta(12)\delta(13)\,, \tag{71}$$

$$P(12) = -i\hbar G_0(12)G_0(21)\,, \tag{72}$$

$$W(12) = v(12) + \iint v(13)P(34)W(42)d3\,d4\,, \tag{73}$$

$$\Sigma(12) = i\hbar G_0(12)W(1^+2)\,, \tag{74}$$

$$G(12) = G_0(12) + \iint G_0(13)\Sigma(34)G(42)d3\,d4\,. \tag{75}$$

The polarization function here corresponds to the bubble diagram of Feynman's diagrammatic approach to many-body perturbation theory and leads to the random-phase approximation for the screened interaction (cf. Fig. 5). The expression for the self-energy in this first iteration coined the name GW approximation.

B Quasiparticle Equation

Inserting (12) into the equation of motion for the Green function of a stationary system in the frequency domain

$$\left[\hbar\omega - \hat{h}_0(\mathbf{r})\right] G(\mathbf{r},\mathbf{r}',\omega) - \int \Sigma(\mathbf{r},\mathbf{r}'';\omega) G(\mathbf{r}'',\mathbf{r}';\omega) d^3r'' = \delta(\mathbf{r}-\mathbf{r}') \;, \tag{76}$$

which is equivalent to the Dyson equation, yields

$$\sum_i \frac{\psi_i^*(\mathbf{r}')}{\hbar\omega - \epsilon_i \mp i\eta} \left\{ \left[\hbar\omega - \hat{h}_0(\mathbf{r})\right] \psi_i(\mathbf{r}) - \int \Sigma(\mathbf{r},\mathbf{r}'';\omega)\, \psi_i(\mathbf{r}'')\, d^3r'' \right\} = \delta(\mathbf{r}-\mathbf{r}') \;. \tag{77}$$

Now we multiply with $(\hbar\omega - \epsilon_j)$ and take the limit $\omega \to \epsilon_j/\hbar$ on both sides. If we assume that the system is nondegenerate (i.e., all ϵ_i are different), the left-hand side becomes

$$\begin{aligned} &\lim_{\omega\to\epsilon_j/\hbar} (\hbar\omega - \epsilon_j) \sum_i \frac{\psi_i^*(\mathbf{r}')}{\hbar\omega - \epsilon_i \mp i\eta} \left\{ \left[\hbar\omega - \hat{h}_0(\mathbf{r})\right] \psi_i(\mathbf{r}) - \int \Sigma(\mathbf{r},\mathbf{r}'';\omega)\, \psi_i(\mathbf{r}'')\, d^3r'' \right\} \\ &\quad = \psi_j^*(\mathbf{r}') \left\{ \left[\epsilon_j - \hat{h}_0(\mathbf{r})\right] \psi_j(\mathbf{r}) - \int \Sigma(\mathbf{r},\mathbf{r}'';\epsilon_j/\hbar)\, \psi_j(\mathbf{r}'')\, d^3r'' \right\} \;, \end{aligned} \tag{78}$$

and the right-hand side becomes

$$\lim_{\omega\to\epsilon_j/\hbar} (\hbar\omega - \epsilon_j)\, \delta(\mathbf{r}-\mathbf{r}') = 0 \,. \tag{79}$$

Since $\psi_j^*(\mathbf{r}')$ does not vanish for all $\mathbf{r}'$, the expression in the curly brackets must be zero. This leads directly to the quasiparticle equation

$$\hat{h}_0(\mathbf{r})\, \psi_j(\mathbf{r}) + \int \Sigma(\mathbf{r},\mathbf{r}'';\epsilon_j/\hbar)\, \psi_j(\mathbf{r}'')\, d^3r'' = \epsilon_j \psi_j(\mathbf{r}) \;. \tag{80}$$

It remains valid in the degenerate case, which is seen as follows. We assume that the solution of Eq. (80) leads to degenerate amplitudes $\psi_j(\mathbf{r})$ and energies ϵ_j. Then we introduce an arbitrary perturbation $\hat{\phi}$, e.g., an additional external potential in $\hat{h}_0(\mathbf{r})$, that breaks the symmetry in such a way that the degeneracy is lifted. For this (nondegenerate) case the above proof applies. The validity of Eq. (80) for the degenerate case is then established by taking the limit $\hat{\phi} \to 0$.

References

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[2] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (Dover, New York, 2003).

[3] G. D. Mahan, *Many-Particle Physics* (Plenum, New York, 1990).

[4] L. Hedin, Phys. Rev. **139**, A796 (1965).

[5] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[6] L. J. Sham and W. Kohn, Phys. Rev. **145**, 561 (1966).

[7] J. P. Perdew, R. G. Parr, M. Levy and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).

[8] W. G. Aulbur, L. Jönsson and J. W. Wilkins, in *Solid State Physics,* edited by H. Ehrenreich and F. Spaepen (Academic, New York, 2000), Vol. 54, p. 1.

[9] P. Duffy, D. P. Chong, M. E. Casida and D. R. Salahub, Phys. Rev. A **50**, 4707 (1994).

[10] W. von der Linden and P. Horsch, Phys. Rev. B **37**, 8351 (1988).

[11] G. E. Engel and B. Farid, Phys. Rev. B **47**, 15931 (1993).

[12] M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **55**, 1418 (1985).

[13] R. W. Godby, M. Schlüter and L. J. Sham, Phys. Rev. Lett. **56**, 2415 (1986).

[14] M. S. Hybertsen and S. G. Louie, Phys. Rev. B **34**, 5390 (1986).

[15] R. W. Godby, M. Schlüter and L. J. Sham, Phys. Rev. B **35**, 4170 (1987).

[16] J. E. Northrup, M. S. Hybertsen and S. G. Louie, Phys. Rev. Lett. **59**, 819 (1987).

[17] R. Keyling, W.-D. Schöne and W. Ekardt, Phys. Rev. B **61**, 1670 (2000).

[18] G. Onida, L. Reining and A. Rubio, Rev. Mod. Phys. **74**, 601 (2002).

[19] M. Petersilka, U. J. Gossmann and E. K. U. Gross, Phys. Rev. Lett. **76**, 1212 (1996).

[20] L. Hedin and S. Lundqvist, in *Solid State Physics*, edited by F. Seitz, D. Turnbull and H. Ehrenreich (Academic, New York, 1969), Vol. 23, p. 1.

A 5 Scattering theory of Bloch wavefunctions off defects in crystals[1]

Phivos Mavropoulos
Institute for Advanced Simulation: Quantum Theory of Materials
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercom-

1 Introduction

Quantum scattering theory is concerned with the problem of a perturbation in a "reference" system with continuous energy spectrum. In the most common form of the theory, the reference system is the system of a beam of freely moving particles in space, while the perturbation is a target that the beam hits and is scattered by. The scattered beam is detected by an apparatus and conclusions on the structure of the target are drawn from the beam distribution in energy, momentum, spin, etc. The central quantity that is measured is the *scattering cross section* and this has to be related by scattering theory to the microscopic structure of the target.

In solid state physics one is often faced with a different kind of problem. The reference system is a crystal with a periodic potential and with electron eigenstates obeying the Bloch theorem. The Bloch states, forming a continuous spectrum given by the band structure, are scattered by imperfections of the periodic potential. These imperfections can be impurity atoms, antisites, self-interstitials, vacancies, etc., or can be excitations of the solid, such as phonons, magnons or excited electrons. Here, there is no prepared beam and detector, while the measured quantity is some material property. Such material properties influenced by scattering can be transport coefficients, e.g., the electrical and heat conductivity, the thermopower, or the spin diffusion length; or they can be equilibrium properties, such as the charge density or spin density in the vicinity of an impurity atom, showing wavy oscillations because of the scattering of electrons off the impurity.

The purpose of this manuscript is to give an introduction to the theory of Bloch electron scattering in solids. We discuss the underlying physical assumptions, introduce the formalism and derive some central results. One of the central quantities derived is the transition rate that can be used in conjunction with the semiclassical Boltzmann transport equation in order to derive transport-related quantities.

We treat the problem of scattering by a localized potential but do not discuss the scattering by a target with internal structure where transitions can occur between the internal degrees of freedom of the scatterer. For example, scattering off an impurity atom or a vacancy can be treated within the formalism that is presented here, but an extension is needed to treat scattering off phonons where one would need to account for exchange of energy and momentum between the phonon and the electron.

In order to keep the manuscript at a reasonable length, the discussion of certain steps in the formalism is avoided. Most important is the omission of the Møller operators that are used for the derivation of some formulas. The presentation here is not mathematically rigorous, since such a task would be a full course on its own right. For a more complete introduction to scattering theory we refer to the literature, especially the books by Bohm [1] and Rodberg and Thaler [2].

2 Green functions and time evolution[2]

Consider an electron under the influence of a potential $V(\boldsymbol{r})$. The Hamiltonian is

$$H = -\nabla^2 + V(\boldsymbol{r}) = H_{\text{free}} + V(\boldsymbol{r}). \tag{1}$$

[2] A large part of this section is adapted from Ref. [3].

Here, and in the following, we use atomic units ($\hbar = 1$, $m_{\rm e} = \frac{1}{2}$, $e = -\sqrt{2}$). The time-dependent Schrödinger equation, determining the time evolution of the electron wavefunction $\psi(t)$, is

$$i\frac{\partial}{\partial t}\psi(t) = H\psi(t) \tag{2}$$

with formal solution

$$\psi(t) = e^{-iHt}\psi(0), \tag{3}$$

and with $\psi(0)$ representing the wavepacket preparation as an initial condition. The quantity $U(t) = e^{-iHt}$ is the time evolution operator in quantum mechanics.
Corresponding to the Schrödinger equation, we define now two propagators, the *retarded Green function* $G^+(t)$ and the *advanced Green function* $G^-(t)$, as

$$\left(i\frac{\partial}{\partial t} - H\right) G^{+,-}(t) = \delta(t) \tag{4}$$

with boundary conditions

$$G^+(t) = 0 \qquad t < 0 \tag{5}$$
$$G^-(t) = 0 \qquad t > 0. \tag{6}$$

The formal solution of these equations is

$$G^+(t) = \begin{cases} -ie^{-iHt} & t > 0 \\ 0 & t < 0 \end{cases} \tag{7}$$

$$G^-(t) = \begin{cases} 0 & t > 0 \\ ie^{-iHt} & t < 0 \end{cases} \tag{8}$$

Evidently, the Green functions coincide, up to a factor, with the time evolution operator; $G^+(t)$ is used to propagate the wavefunction forward in time, and $G^-(t)$ to propagate the wavefunction backward in time:

$$\psi(t) = iG^+(t-t')\psi(t') \qquad t' < t \tag{9}$$
$$\psi(t) = -iG^-(t-t')\psi(t') \qquad t < t'. \tag{10}$$

In the following, unless explicitly stated, we shall restrict the discussion to the retarded Green function G^+ and drop the index $+$.
Given a perturbing potential $\Delta V(\boldsymbol{r})$ added to a Hamiltonian $\mathring{H}$, it follows from equation (4) that the Green function G corresponding to the new Hamiltonian $H = \mathring{H} + \Delta V$ is related to the Green function $\mathring{G}$ corresponding to $\mathring{H}$ via the Dyson integral equation

$$G(t) = \mathring{G}(t) + \int_0^t \mathring{G}(t-t')\ \Delta V\, G(t')\, dt'. \tag{11}$$

Furthermore, it can be proven that an incoming wavepacket $\mathring{\psi}$, which without the interaction with ΔV would evolve as $\mathring{\psi}(t)$, evolves into the wavefunction $\psi(t)$ as

$$\psi(t) = \mathring{\psi}(t) + \int_{-\infty}^t \mathring{G}(t-t')\ \Delta V\, \psi(t')\, dt' \tag{12}$$
$$= \mathring{\psi}(t) + \int_{-\infty}^t G(t-t')\ \Delta V\, \mathring{\psi}(t')\, dt' \tag{13}$$

These are two equivalent forms of the *Lippmann-Schwinger equation*; in the former, Eq. (12), the reference Green function enters and the solution ψ appears on both sides of the equation; in the latter, Eq. (13), the Green function of the full system enters and the solution appears only on the left-hand side.

2.1 Energy-dependent Green functions

A Fourier transform of the Green function,

$$G(E) = \int_{-\infty}^{\infty} G(t)\, e^{i(E+i\eta)t}\, dt \tag{14}$$

results in the formal solution

$$G(E) = (E + i\eta - H)^{-1}. \tag{15}$$

Here, η is a positive real number which ensures convergence of the integral in (14) for $t \to \infty$ and is to be taken in the end to the limit $\eta \to 0$; this limit is implied by the notation $E \pm i0$ that we will use later. More generally, one may define the (time-independent) Green function as the resolvent of the (time-independent) Schrödinger equation, via the operator equation

$$G(z) = (z - H)^{-1} \tag{16}$$

for an arbitrary complex energy z (as long as $(z-H)$ can be inverted). The singularities of $G(z)$ along the real axis $z = E$ determine the eigenvalue spectrum; in particular, $G(E)$ has poles at the eigenenergies of the bound states and a branch cut along the energies of the continuous spectrum. For $\mathrm{Im}z > 0$, $G(z)$ is an analytical function of z. In terms of a complete set of eigenfunctions of H, $|\psi_i\rangle$, corresponding to eigenvalues ϵ_i, the following spectral representation can be obtained:

$$G(z) = \sum_i \frac{|\psi_i\rangle\langle\psi_i|}{z - \epsilon_i}. \tag{17}$$

Represented in real space (and ignoring here for simplicity the spin variables), the above equation becomes

$$G(\boldsymbol{r}, \boldsymbol{r}'; z) = \sum_i \frac{\psi_i(\boldsymbol{r})\, \psi_i^*(\boldsymbol{r}')}{z - \epsilon_i} \tag{18}$$

representing, in the limit of $\mathrm{Im}z = \eta \to 0^+$, an outgoing wave at $\boldsymbol{r}$ with a source term at $\boldsymbol{r}'$. From the above equation one can see that the imaginary part of G is directly related to the spectrally- and space-resolved density of states $n(\boldsymbol{r}; E)$ (for real E):

$$n(\boldsymbol{r}; E) = -\frac{1}{\pi}\mathrm{Im}\, G(\boldsymbol{r}, \boldsymbol{r}; E). \tag{19}$$

This follows from the Dirac identity,

$$\int_{-\infty}^{\infty} \frac{f(x)}{x - x_0 \pm i\eta} dx = P\left(\int_{-\infty}^{\infty} \frac{f(x)}{x - x_0} dx\right) \mp i\pi f(x_0) \tag{20}$$

where P stands for the Cauchy principal part of the integral. One can deduce from (19) an expression for the spectral density of states,

$$n(E) = -\frac{1}{\pi}\mathrm{Im} \int G(\boldsymbol{r}, \boldsymbol{r}; E)\, d^3r = -\frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\, G(E), \tag{21}$$

where the last step stresses that the trace of the operator $G(E)$ can be taken in any basis, not just in real-space representation. On the other hand, it can be seen from Eq. (19) after integrating over all occupied states that the charge density is found as an integral of $n(\boldsymbol{r};E)$ over the energies convoluted with a Fermi function[3] $f_0(E;T) = 1/[1+\exp(\frac{E-E_\mathrm{F}}{k_\mathrm{B}T})]$:

$$\rho(\boldsymbol{r}) = -\frac{1}{\pi}\mathrm{Im}\int_{-\infty}^{\infty} G(\boldsymbol{r},\boldsymbol{r};E)\, f_T(E)\, dE = -\frac{1}{\pi}\mathrm{Im}\int_{-\infty}^{\infty} \mathrm{Tr}\left[\mathrm{Pr}_{\hat{\boldsymbol{r}}}\, G(E)\right]\, f_0(E;T)\, dE \quad (22)$$

where $\mathrm{Pr}_{\hat{\boldsymbol{r}}} = |\boldsymbol{r}\rangle\delta(\boldsymbol{r}-\boldsymbol{r}')\langle\boldsymbol{r}'|$ is the projection operator on the position $\boldsymbol{r}$ and the trace implies an integration over $\boldsymbol{r}$ and $\boldsymbol{r}'$, i.e., $\mathrm{Tr}\left[\mathrm{Pr}_{\hat{\boldsymbol{r}}}\, G(E)\right] = \int d^3r\, d^3r'\, \delta(\boldsymbol{r}-\boldsymbol{r}')G(\boldsymbol{r}',\boldsymbol{r};E)$. If a spin degree of freedom is present, then the Green function is a matrix in spin space and the trace must be taken also over spins.
In general, from the spectral representation (17) it follows that the expectation value (averaged over all occupied electron states) of any physical quantity, represented by an operator $\hat{A}$, can be harvested via the relation

$$\langle\hat{A}\rangle = -\frac{1}{\pi}\mathrm{Im}\int_{-\infty}^{\infty} \mathrm{Tr}\left[\hat{A}\, G(E)\right]\, f_T(E)\, dE. \quad (23)$$

Therefore, the Green function contains all information which is given by the eigenfunctions. If the Green function can be computed, then all physical properties of the system can be found.

2.2 Relation between perturbed and unperturbed system; the Dyson equation

The time-dependent Dyson equation (11) has its energy-dependent counterpart. Given from (16) that for the reference system $\mathring{G}^{-1}(E) = E - \mathring{H}$, while for the perturbed system $G^{-1}(E) = E - (\mathring{H} + \Delta V)$, we conclude that the two Green functions are connected via

$$G^{-1}(E) = \mathring{G}^{-1}(E) - \Delta V. \quad (24)$$

The Dyson equation can be rewritten in the following forms (which can be directly verified by substitution in (24)):

$$\begin{aligned} G(E) &= \left[1 - \mathring{G}(E)\ \Delta V\right]^{-1} \mathring{G}(E) && (25)\\ &= \mathring{G}(E)\left[1 - \Delta V\, \mathring{G}(E)\right]^{-1} && (26)\\ &= \mathring{G}(E) + \mathring{G}(E)\ \Delta V\, G(E). && (27) \end{aligned}$$

The last expression, although not in a closed form, reminds of expression (11) and allows for an interpretation of the Dyson equation via multiple-scattering events by the perturbing potential ΔV if we expand G on the right-hand side (rhs):

$$G(E) = \mathring{G}(E) + \mathring{G}(E)\ \Delta V\, \mathring{G}(E) + \mathring{G}(E)\ \Delta V \mathring{G}(E)\ \Delta V\, \mathring{G}(E) + \cdots; \quad (28)$$

this is the analogue of a Born series expansion for the Green function.

[3] This is of course valid under the assumption that we have non-interacting electrons.

Also the Lippmann-Schwinger equation (12) has a time-independent counterpart. Observing that the Schrödinger equation for the perturbed system can be written as

$$(E - H)\,|\psi\rangle = \Delta V\,|\psi\rangle, \tag{29}$$

we can verify by substitution that the solution $|\psi\rangle$ can be written in terms of the unperturbed eigenstates $|\mathring{\psi}\rangle$ as

$$|\psi\rangle = |\mathring{\psi}\rangle + \mathring{G}(E)\ \Delta V\,|\psi\rangle \tag{30}$$

where, loosely speaking, the first term on the rhs represents the incoming wave and the second term the scattered wave, with the interference of the two producing the total wave $|\psi\rangle$. Eq. (30) has the peculiarity that the solution $|\psi\rangle$ occurs on both sides of the equation. Expanding $|\psi\rangle$ on the rhs of this Lippmann-Schwinger equation leads to the Born series

$$|\psi\rangle = |\mathring{\psi}\rangle + \mathring{G}(E)\ \Delta V\,|\mathring{\psi}\rangle + \mathring{G}(E)\ \Delta V\,\mathring{G}(E)\ \Delta V\,|\mathring{\psi}\rangle + \cdots . \tag{31}$$

The Lippmann-Schwinger equation can be also written in terms of the perturbed Green function $G(E)$ instead of the unperturbed $\mathring{G}(E)$ as:

$$|\psi\rangle = |\mathring{\psi}\rangle + G(E)\ \Delta V\,|\mathring{\psi}\rangle. \tag{32}$$

where the unperturbed solution, $|\mathring{\psi}\rangle$, enters in the rhs instead of the perturbed $|\psi\rangle$ that enters in the rhs of Eq. (32). The equivalence between Eq. (30) and Eq. (32) can be seen if one inserts the expansion Eq. (28) in Eq. (32) and compares with Eq. (31); alternatively one can formally write the solution of Eq. (30) as $|\psi\rangle = [1 - \mathring{G}(E)\,\Delta V]^{-1}|\mathring{\psi}\rangle$ and observe that $[1 - \mathring{G}(E)\,\Delta V]^{-1} = [1 + G(E)\,\Delta V]$ by virtue of the Dyson equation (27). The form Eq. (32) of the Lippmann-Schwinger equation is used in practice when $G(E)$ has already been found by means of the Dyson equation.

In the case that E does not belong to the spectrum of the unperturbed Hamiltonian H, $|\mathring{\psi}\rangle = 0$ and from Eq. (30) we obtain

$$|\psi\rangle = \mathring{G}(E)\ \Delta V\,|\psi\rangle. \tag{33}$$

This expression can be used to obtain the discrete spectrum of H by e.g. an expansion in a basis set $\{|n\rangle\}$, $|\psi\rangle = \sum_n c_n|n\rangle$, leading to the homogeneous system of algebraic linear equations $\sum_{n'}(\delta_{nn'} - [\mathring{G}(E)\Delta V]_{nn'})c_{n'}(E) = 0$. The real-valued energies E where the system has a non-trivial solution form the discrete spectrum of the system. In this way one can e.g. locate the bound states introduced by defects in the band gap of semiconductors.

The Dyson equation provides a straightforward way for the calculation of response functions. Consider a small perturbation $\delta V = \alpha\,\hat{V}_\mathrm{p}$, where $\hat{V}_\mathrm{p}$ is a hermitian operator representing the perturbing potential and α is a small parameter controlling the strength of the perturbation. We seek the response, in the linear regime, of an observable $\hat{A}$, $\delta\langle\hat{A}\rangle := \langle\hat{A}\rangle_\alpha - \langle\hat{A}\rangle_0$, where $\langle\hat{A}\rangle_\alpha$ is the expectation value of $\hat{A}$ in the presence of the perturbation and $\langle\hat{A}\rangle_0$ is its expectation value without the perturbation. For example, $\hat{V}_\mathrm{p}$ could be a magnetic field acting at some position $\boldsymbol{r}'$ and $\hat{A}$ could be the change in magnetization at another position $\boldsymbol{r}$, so that the sought-after quantity is the magnetic susceptibility tensor $\chi(\boldsymbol{r}, \boldsymbol{r}')$ with $\delta\boldsymbol{m}(\boldsymbol{r}) = \chi(\boldsymbol{r}, \boldsymbol{r}')\delta\boldsymbol{B}(\boldsymbol{r}')$. Since the perturbation is small, it is convenient to use the expansion (28). From Eq. (23) (setting for simplicity $T = 0$) we have

$$\begin{aligned}\langle\hat{A}\rangle_\alpha &= -\frac{1}{\pi}\mathrm{Im}\int^{E_\mathrm{F}} \mathrm{Tr}\left[\hat{A}\,\mathring{G}(E)\right] dE - \alpha\frac{1}{\pi}\mathrm{Im}\int^{E_\mathrm{F}} \mathrm{Tr}\left[\hat{A}\,\mathring{G}(E)\,\hat{V}_\mathrm{p}\,\mathring{G}(E)\right] dE \\ &= \langle\hat{A}\rangle_0 - \alpha\frac{1}{\pi}\mathrm{Im}\int^{E_\mathrm{F}} \mathrm{Tr}\left[\hat{A}\,\mathring{G}(E)\,\hat{V}_\mathrm{p}\,\mathring{G}(E)\right] dE,\end{aligned} \tag{34}$$

i.e., we find that

$$\left.\frac{d\langle\hat{A}\rangle}{d\alpha}\right|_{\alpha=0} = -\frac{1}{\pi}\mathrm{Im}\int^{E_\mathrm{F}} \mathrm{Tr}\left[\hat{A}\,\mathring{G}(E)\,\hat{V}_\mathrm{p}\,\mathring{G}(E)\right] dE. \tag{35}$$

As an application we consider the magnetic susceptibility. The observed quantity is then the magnetization $\boldsymbol{m}(\boldsymbol{r}) = \langle\boldsymbol{\sigma}\ \mathrm{Pr}_{\boldsymbol{r}}\rangle$, i.e., $\hat{A} = \boldsymbol{\sigma}\ \mathrm{Pr}_{\boldsymbol{r}} = \boldsymbol{\sigma}\,|\boldsymbol{r}\rangle\delta(\boldsymbol{r}-\boldsymbol{r}_1)\langle\boldsymbol{r}_1|$, while the perturbation is a weak magnetic field at $\boldsymbol{r}'$ coupled to the electron system via the operator[4] $\alpha\hat{V}_\mathrm{p} = \boldsymbol{B}(\boldsymbol{r}')\cdot\boldsymbol{\sigma}\ \mathrm{Pr}_{\boldsymbol{r}'} = \boldsymbol{B}(\boldsymbol{r}')\cdot\boldsymbol{\sigma}\,|\boldsymbol{r}'\rangle\delta(\boldsymbol{r}'-\boldsymbol{r}_2)\langle\boldsymbol{r}_2|$, where we have written explicitly the dependence on two variables, $(\boldsymbol{r},\boldsymbol{r}_1)$ and also $(\boldsymbol{r}',\boldsymbol{r}_2)$, to stress the matrix-form of the two operators. Applying Eq. (35) we obtain for the i^th component of the magnetization and the j^th component of the field ($i,j\in\{x,y,z\}$):

$$m_i(\boldsymbol{r}) = -\frac{1}{\pi}\mathrm{Im}\int^{E_\mathrm{F}} \mathrm{Tr}_s\left[\sigma_i\,\mathring{G}(\boldsymbol{r},\boldsymbol{r}';E)\,\sigma_j\,\mathring{G}(\boldsymbol{r}',\boldsymbol{r};E)\right] dE\ B_j(\boldsymbol{r}'). \tag{36}$$

with Tr_s the trace over the spin-degrees of freedom (the Green function $\mathring{G}$ is a 2×2 matrix in spin space) while the trace over the spatial degrees of freedom, $(\boldsymbol{r},\boldsymbol{r}_1)$ and $(\boldsymbol{r}',\boldsymbol{r}_2)$, was taken through respective integrations d^3r_1 and d^3r_2 that were lifted by the δ-functions. From Eq. (36) we read out the susceptibility tensor

$$\chi_{ij}(\boldsymbol{r},\boldsymbol{r}') = -\frac{1}{\pi}\mathrm{Im}\int^{E_\mathrm{F}} \mathrm{Tr}_s\left[\sigma_i\,\mathring{G}(\boldsymbol{r},\boldsymbol{r}';E)\,\sigma_j\,\mathring{G}(\boldsymbol{r}',\boldsymbol{r};E)\right] dE \tag{37}$$

3 $\mathcal{S}$-matrix and $\mathcal{T}$-matrix

We proceed here by introducing a fundamental quantity of scattering theory, the $\mathcal{S}$-matrix. It contains the full information on the scattering properties of a system.

We start by introducing some notation details. We consider the scattering of Bloch electrons in a crystal lattice off localized defects, e.g. impurity atoms, in the approximation of one-electron theory, i.e., assuming that we have a single-electron wavefunction traveling in the crystal potential $\mathring{V}$. The defects produce a perturbation of the crystal potential, $\Delta V = V - \mathring{V}$, that has a finite range including the perturbation of the neighbouring atoms. We denote the Bloch wavefunctions as $|n\boldsymbol{k}\rangle$ where $\boldsymbol{k}$ is the crystal momentum and n is a band index including, whenever necessary, quantum numbers that correspond to some degeneracy (e.g. the spin degeneracy in non-magnetic systems).

In view of the need to chart constant-energy surfaces in reciprocal space (e.g. the Fermi surface), defined by the relation $E_{n\boldsymbol{k}} = E$, we introduce a local coordinate $(\boldsymbol{k}_\perp,\boldsymbol{k}_\parallel)$ on the surface with a decomposition of the crystal momentum element $d^3k = d\boldsymbol{k}_\perp d\boldsymbol{k}_\parallel$, where $\boldsymbol{k}_\perp$ is normal to the constant-energy surface and $\boldsymbol{k}_\parallel$ is oriented in the surface. Thus $\boldsymbol{k}_\perp$ is in the direction of the velocity,[5] $\boldsymbol{v}_{n\boldsymbol{k}} := \boldsymbol{\nabla}_{\boldsymbol{k}}E_{n\boldsymbol{k}}$. Using this notation the normalization between Bloch functions reads

$$\langle n'\boldsymbol{k}'|n\boldsymbol{k}\rangle = \Omega_\mathrm{rec}\,\delta_{nn'}\delta(\boldsymbol{k}-\boldsymbol{k}') = \Omega_\mathrm{rec}\,\delta_{nn'}\,v_{n\boldsymbol{k}}\,\delta(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}'})\,\delta(\boldsymbol{k}_\parallel-\boldsymbol{k}'_\parallel) \tag{38}$$

where $\Omega_\mathrm{rec} = \Omega_\mathrm{BZ}/N_\mathrm{cr} = (2\pi)^3/\Omega_\mathrm{cryst}$ is the volume element of the reciprocal lattice with Ω_BZ is the Brillouin zone volume, Ω_cryst the crystal volume and and N_cr the number of primitive

[4] The parameter α is absorbed in $\boldsymbol{B}(\boldsymbol{r}')$.

[5] We define the velocity in this way having in mind that, in the presence of spin-orbit coupling, there is an additional "anomalous" contribution to the true velocity of a wavepacket.

cells in the crystal. We denote the integration weight in the Brillouin zone by $d\boldsymbol{k} := \Omega_{\rm rec}^{-1} d^3k$ (with obvious adjustments in two dimensional systems such as crystalline thin films).
In a scattering process we consider a wavepacket that approaches a region that contains a potential perturbation ΔV. In the beginning, at times long before the scattering takes place ($t \to -\infty$), the wavepacket is evolving in time according to the unperturbed Hamiltonian $\mathring{H}$; later it reaches the point of scattering where $\Delta V \neq 0$, e.g. a defect in the crystal, and evolves according to the perturbed Hamiltonian $H = \mathring{H} + \Delta V$; and at much later times ($t \to \infty$) it has left the region where $\Delta V \neq 0$ and propagates again according to $\mathring{H}$.
The mapping from the initial to the final form of the wavepacket is done by a unitary operator, the $\mathcal{S}$ operator, that is represented in the basis of eigenstates of $\mathring{H}$ by the $\mathcal{S}$-matrix. For a definition of the $\mathcal{S}$ operator the limits $t \to \pm\infty$ must be properly taken because of the oscillating phase $\exp(-iEt)$ of the wavefunction. This is conveniently done, e.g. in the interaction picture of quantum mechanics, or by taking a sharp distribution around an energy E and using the scattering equations in the energy domain. We follow the latter course here.

3.1 In-states and out-states; $\mathcal{S}$-matrix

Far away from the defect and at early times, before the scattering takes place, a wavepacket $\psi(t)$ with a finite but sharp distribution around the crystal momentum $\boldsymbol{k}$ at band n propagates according to the crystal Hamiltonian:

$$\psi(t) = \int d\boldsymbol{q} A(\boldsymbol{q}-\boldsymbol{k}) |n\boldsymbol{q}\rangle e^{-iE_{n\boldsymbol{q}}t}, \quad t \to -\infty \tag{39}$$

where $A(\boldsymbol{q}-\boldsymbol{k})$ is the amplitude of the wave packet at each $\boldsymbol{q}$ close to $\boldsymbol{k}$ (in the limit of a very sharp distribution around $\boldsymbol{k}$ the squared amplitude $|A(\boldsymbol{q}-\boldsymbol{k})|^2$ will become a $\delta((\boldsymbol{q}-\boldsymbol{k}))$ function). During and after the scattering takes place, the wave packet becomes

$$\psi(t) = \int d\boldsymbol{k}' A(\boldsymbol{k}'-\boldsymbol{k}) |n\boldsymbol{k}';+\rangle e^{-iE_{n\boldsymbol{k}'}t} \tag{40}$$

where $|n,\boldsymbol{k};+\rangle$ are related to the Bloch waves $|n,\boldsymbol{k}\rangle$ according to the Lipmann-Schwinger equation Eq. (32) which in the present Bloch-wave notation reads

$$\begin{aligned} |n,\boldsymbol{k};+\rangle &= |n\boldsymbol{k}\rangle + G^+(E_{n\boldsymbol{k}})\,\Delta V |n\boldsymbol{k}\rangle = |n\boldsymbol{k}\rangle + [E_{n\boldsymbol{k}} - H + i\eta]^{-1}\,\Delta V |n\boldsymbol{k}\rangle & (41)\\ &= |n\boldsymbol{k}\rangle + \mathring{G}^+(E_{n\boldsymbol{k}})\,\Delta V |n\boldsymbol{k};+\rangle = |n\boldsymbol{k}\rangle + [E_{n\boldsymbol{k}} - \mathring{H} + i\eta]^{-1}\,\Delta V |n\boldsymbol{k}\rangle & (42)\end{aligned}$$

where $G^+(E_{n\boldsymbol{k}})$ is the retarded Green function of the perturbed system (crystal plus defect) and ΔV the perturbation. Here, η denotes an infinitesimally small positive imaginary part of the energy that is taken to zero at the end, as in Eq. (15), which basically ensures that we are working with the retarded Green function. In the limit of early times, $t \to -\infty$, the term $\int d\boldsymbol{k}' [E_{n\boldsymbol{k}'} - H + i\eta]^{-1}\,\Delta V |n\boldsymbol{k}'\rangle e^{-iE_{n\boldsymbol{k}'}t}$ vanishes, as we show in the Appendix. Eq. (42) is the equivalent form to Eq. (30) where the crystal Green function $\mathring{G}^+$ and the perturbed wavefunction appear on the rhs.
Equations (39-40) motivate us to define the set of in-states as all possible states of the form $|n\boldsymbol{k}\rangle$ that propagate in the unperturbed system according to the unperturbed crystal Hamiltonian and that are incoming towards the defect. Thus $|n\boldsymbol{k};+\rangle$ are states that in the remote past were in-states but at the end have a complex form as superpositions of Bloch waves. The indices $n\boldsymbol{k}$ here do not mean that $|n\boldsymbol{k};+\rangle$ is a state of crystal momentum $\boldsymbol{k}$ at band n (it is not), but only

that it originates from such a states in the remote past; i.e., $n\boldsymbol{k}$ indicate the boundary condition of the incoming wave.
In an analogous way we define the set of out-states as all possible states that have the form of Bloch waves $|n,\boldsymbol{k}\rangle$ and propagate according to the unperturbed crystal Hamiltonian in the distant future. These are the result of propagation and interaction with the defect of states that we symbolize as $|n\boldsymbol{k};-\rangle$, that had a complex form in the past, before and during the scattering, the form being determined by the requirement that at $t\to\infty$ the superposition $|n\boldsymbol{k};-\rangle$ evolves into a single Bloch wave $|n\boldsymbol{k}\rangle$. Again, the indices $n\boldsymbol{k}$ in $|n\boldsymbol{k};-\rangle$ do not mean that the state is a single Bloch wave, but that it evolves into one in the distant future. The states $|n\boldsymbol{k};-\rangle$ are again the solution of the Lippmann-Schwinger equation, but with the advanced Green function:

$$\begin{aligned} |n\boldsymbol{k};-\rangle &= |n\boldsymbol{k}\rangle + G^{-}(E_{n\boldsymbol{k}})\Delta V|n\boldsymbol{k}\rangle = |n\boldsymbol{k}\rangle + (E_{n\boldsymbol{k}} - H - i0)^{-1}\,\Delta V|n\boldsymbol{k}\rangle \qquad (43)\\ &= |n\boldsymbol{k}\rangle + \mathring{G}^{-}(E_{n\boldsymbol{k}})\Delta V|n\boldsymbol{k}\rangle = |n\boldsymbol{k}\rangle + (E_{n\boldsymbol{k}} - \mathring{H} - i0)^{-1}\,\Delta V|n\boldsymbol{k};-\rangle. \qquad (44)\end{aligned}$$

The Hilbert spaces spanned by the in and out states in a crystal with a localized defect coincide with each other and with the set of Bloch states, but the distinction is made because there can occur cases in scattering theory when they in- and out-spaces are different. One such example is a system of sample attached to two wires, where the in-states are the states of the leads propagating towards the sample while the out-states are the states propagating away from the wire.
The normalization properties of the states $|n\boldsymbol{k};\pm\rangle$ is the same as the one of the unperturbed Bloch states Eq. (38) ([1], Chap. XIV.2 and XV.A):

$$\begin{aligned} \langle n'\boldsymbol{k}';+|n\boldsymbol{k};+\rangle = \langle n'\boldsymbol{k}';-|n\boldsymbol{k};-\rangle &= \langle n'\boldsymbol{k}'|n\boldsymbol{k}\rangle\\ &= \delta_{n'n}\delta_{\boldsymbol{k}'\boldsymbol{k}}\\ &= \Omega_{\rm rec}\, v_{n\boldsymbol{k}}\,\delta_{n'n}\,\delta(\boldsymbol{k}'_{\parallel}-\boldsymbol{k}_{\parallel})\,\delta(E_{n'\boldsymbol{k}'}-E_{n\boldsymbol{k}}) \qquad (45)\end{aligned}$$

Using the identity that relates the Green function to the spectral function,

$$(E-H+i0)^{-1} - (E-H-i0)^{-1} = -2\pi i\delta(E-H) \qquad (46)$$

we also obtain a relation between $|n\boldsymbol{k};+\rangle$ and $|n\boldsymbol{k};-\rangle$:

$$|n\boldsymbol{k};+\rangle - |n\boldsymbol{k};-\rangle = -2\pi i\delta(E_{n\boldsymbol{k}}-H)\,\Delta V|n\boldsymbol{k}\rangle \qquad (47)$$

that we will use later.
We are interested in the amplitude that an in-state $|n\boldsymbol{k}\rangle$ at $t\to-\infty$ evolves into an out-state $|n'\boldsymbol{k}'\rangle$ at $t\to\infty$. The operator mapping in-states to out-states is the $\mathcal{S}$-operator and its representation is given by the $\mathcal{S}$-matrix with elements

$$\mathcal{S}_{n'\boldsymbol{k}';n\boldsymbol{k}} := \langle n'\boldsymbol{k}'|\mathcal{S}|n\boldsymbol{k}\rangle = \langle n',\boldsymbol{k}';-|n,\boldsymbol{k};+\rangle. \qquad (48)$$

This is one of more possible equivalent definitions of the $\mathcal{S}$-matrix. The $\mathcal{S}$-matrix is unitary, $\mathcal{S}^{-1}=\mathcal{S}^{\dagger}$, as required by the conservation of probability. One can thus view the $\mathcal{S}$ as a transformation between the in and out Hilbert spaces.
The second part of Eq. (48) can be shown in the folllowing way. One considers the wavepacket corresponding to the in-state, Eq. (40), and the analogous wavepacket corresponding to the final, out-state, $\psi'(t) = \int d\boldsymbol{q}' A'(\boldsymbol{q}'-\boldsymbol{k}')|n'\boldsymbol{q}';-\rangle e^{-iE_{n'\boldsymbol{q}'}t}$, where the distributions A and A' are peaked around certain $\boldsymbol{k}$ and $\boldsymbol{k}'$. The scalar product $\big(\psi'(t),\psi(t)\big)$ represents the amplitude

that a state prepared in the form (39) in the remote past, peaked around $n\boldsymbol{k}$, is measured to be the state $\psi'(t)$ at time t. Taking $t \to \infty$, the state ψ' is nothing else but the out-state $|n'\boldsymbol{k}'\rangle$ (in the limit that the distribution A' becomes a delta-function around $\boldsymbol{k}'$), i.e., the desired $\mathcal{S}$-matrix element. But $\big(\psi'(t), \psi(t)\big) = \int d\boldsymbol{q}' \int d\boldsymbol{q} A^*(\boldsymbol{q}' - \boldsymbol{k}')A(\boldsymbol{q} - \boldsymbol{k})\langle n'\boldsymbol{q}'; -|n\boldsymbol{q}; +\rangle e^{-i(E_{n\boldsymbol{q}} - E_{n'\boldsymbol{q}'})t}$. It can be shown ([1] Chap. XV.3) that this scalar product is independent of time, thus evaluating it at any time, e.g. $t = 0$, and taking the limit of A and A' to become δ-functions at the end, we obtain $\big(\psi'(t), \psi(t)\big) = \langle n'\boldsymbol{k}'; -|n\boldsymbol{k}; +\rangle$, which is the expression in Eq. (48). In the following section we will see that the $\mathcal{S}$-matrix is energy-conserving.

3.2 $\mathcal{T}$-matrix

The amount of $\mathcal{S}$-matrix elements is larger than one is interested in in a scattering problem, containing information about scattered but also unscattered states. It is convenient to restrict the phase space of the problem to only the scattered states. This is done by introducing the $\mathcal{T}$-matrix. In order to introduce it we proceed by formally calculating the $\mathcal{S}$-element according to (48):

$$\begin{aligned}
\mathcal{S}_{n'\boldsymbol{k}';n\boldsymbol{k}} &= \langle n'\boldsymbol{k}'; -|n\boldsymbol{k}; +\rangle \\
&= \big[\langle n'\boldsymbol{k}'; +| - 2\pi i \langle n'\boldsymbol{k}'|\,\Delta V \delta(E_{n'\boldsymbol{k}'} - H)\,\big]\,|n\boldsymbol{k}; +\rangle \\
&= \langle n'\boldsymbol{k}'; +|n\boldsymbol{k}; +\rangle - 2\pi i \delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}})\,\langle n'\boldsymbol{k}'|\,\Delta V|n\boldsymbol{k}; +\rangle \\
&= \delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}})\big[\Omega_{\rm rec}\, v_{n\boldsymbol{k}}\, \delta(\boldsymbol{k}'_{\parallel} - \boldsymbol{k}_{\parallel})\delta_{n'n} - 2\pi i \langle n'\boldsymbol{k}'|\,\Delta V|n\boldsymbol{k}; +\rangle\big] \\
&=: \delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}})\big[\Omega_{\rm rec}\; v_{n\boldsymbol{k}}\, \delta(\boldsymbol{k}'_{\parallel} - \boldsymbol{k}_{\parallel})\delta_{n'n} - 2\pi i\, \mathcal{T}_{n'\boldsymbol{k}';n\boldsymbol{k}}\big] \qquad (49)
\end{aligned}$$

Here, in the first step we used Eq. (47), in the third step we used the ortho-normalization Eq. (45) while the final step defines the $\mathcal{T}$-matrix element between states $|n\boldsymbol{k}\rangle$ and $|n'\boldsymbol{k}'\rangle$ as

$$\begin{aligned}
\mathcal{T}_{n'\boldsymbol{k}';n\boldsymbol{k}} &:= \langle n'\boldsymbol{k}'|\,\Delta V|n\boldsymbol{k}; +\rangle \\
&= \langle n'\boldsymbol{k}'|\big[\,\Delta V + \,\Delta V G^+(E_{n\boldsymbol{k}})\,\Delta V\big]|n\boldsymbol{k}\rangle \qquad (50)
\end{aligned}$$

where Eq. (41) was used. From the δ-function in Eq. (49) it is clear that the $\mathcal{S}$-matrix is energy conserving which is expected since it is basically equivlent to the time-evolution operator $U_{\rm I}(\infty, -\infty)$ in the interaction picture (we have assumed from the outset that the crystal Hamiltonian and the perturbing potential ΔV are time-independent). However, while the $\mathcal{S}$-matrix elements vanish for $E_{n\boldsymbol{k}} \neq E_{n'\boldsymbol{k}'}$, the same is not necessarily true for $\mathcal{T}$-matrix elements. This poses no inconsistency with energy conservation, because only the on-energy-shell part of the $\mathcal{T}$-matrix ($E_{n\boldsymbol{k}} = E_{n'\boldsymbol{k}'}$) leads to observable quantities; the off-shell part vanishes when convoluted with $\delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}})$ in Eq. (49). Formally, Eq. (49) can be written in the more compact form

$$\mathcal{S} = 1 - 2\pi i \delta(E - \mathring{H})\mathcal{T}. \qquad (51)$$

with E the energy of the initial state.

In Eq. (49) we chose to eliminate $\langle n'\boldsymbol{k}'; -|$ by virtue of Eq. (47). We could equally well have eliminated $|n\boldsymbol{k}; +\rangle$, that would lead to Eq. (49) but with a different definition of the $\mathcal{T}$-matrix, $\mathcal{T}^-_{n'\boldsymbol{k}';n\boldsymbol{k}} = \langle n'\boldsymbol{k}'; -|\Delta V|n\boldsymbol{k}\rangle = \langle n'\boldsymbol{k}'|\big[\,\Delta V + \,\Delta V G^-(E_{n\boldsymbol{k}})\Delta V\big]|n\boldsymbol{k}\rangle$, i.e., with the advanced instead of the retarded Green function. The two definitions give identical matrix elements on the energy shell (because they are both derived from $\langle n'\boldsymbol{k}'; -|n\boldsymbol{k}; +\rangle$), i.e., where these are

physically meaningful. Thus we retain Eq. (50) as the definition of the $\mathcal{T}$-matrix and define more generally the $\mathcal{T}$-operator as

$$\mathcal{T}(E) = \Delta V + \Delta V G^{+}(E)\,\Delta V \tag{52}$$

that, for practical purposes, can be conveniently evaluated as a finite-sized matrix in any localized basis set with truncation of the number of local orbitals per atom. This follows from the assumption that the perturbation ΔV is localized in space around the defect and from the fact that ΔV multiplies the non-localized Green function $G^{+}(E)$ both from the right and from the left. The real-space representation of $\mathcal{T}$ (assuming a local potential) is:

$$\mathcal{T}(\boldsymbol{r},\boldsymbol{r}';E) = \Delta V(\boldsymbol{r})\delta(\boldsymbol{r}-\boldsymbol{r}') + \Delta V(\boldsymbol{r})G^{+}(\boldsymbol{r},\boldsymbol{r}';E)\,\Delta V(\boldsymbol{r}') \tag{53}$$

that obviously vanishes wherever $\Delta V(\boldsymbol{r}) = 0$. The most usual form of non-local potential, the spin-orbit coupling, is also approximated by a term that is localized on each atomic site with zero inter-site contributions, again confining $\mathcal{T}$ in a finite number of sites.
As we will discuss later, knowledge of the $\mathcal{T}$-matrix allows for a calculation of transition-probability rates between states leading to a number of derived measurable properties in solids such as relaxation times or transport properties. Also, the $\mathcal{T}$-matrix is related to the integrated density of states leading e.g. to the Friedel sum rule connecting the scattering phase shifts to the total charge displaced by a defect in a crystal.

3.3 Expansion of the $\mathcal{T}$-matrix; multiple scattering

Equation (52) together with Eq. (27) and Eq. (28) for the Green function suggest the following expansion of the $\mathcal{T}$-matrix:

$$\begin{aligned}
\mathcal{T}(E) &= \Delta V + \Delta V \mathring{G}(E)\,\mathcal{T}(E) && (54)\\
&= \Delta V[1 - \mathring{G}(E)\,\Delta V]^{-1} && (55)\\
&= \Delta V + \Delta V \mathring{G}(E)\,\Delta V + \Delta V \mathring{G}(E)\,\Delta V \mathring{G}(E)\,\Delta V + \cdots && (56)
\end{aligned}$$

Here we have simplified the notation by accepting that we refer to advanced Green function omitting the superscript "+". Keeping the 1st, 2nd, etc., term in Eq. (56) corresponds to treating the scattered state $|n\boldsymbol{k};+\rangle$ to the 1st, 2nd, etc., order in the potential in the Born expansion series of the Lippmann-Schwinger equation Eq. (31). This expansion proves useful if the potential is weak enough to justify a truncation at some low order; if the expansion is not convergent (which can be the case e.g. if the potential is too strong or if the energy is near a scattering resonance causing $\mathcal{T}$ to vary too fast with energy), then Eq. (56) has no meaning, while Eq. (52) and Eq. (54) are still valid. However, the concept of the Born expansion paves the way to a perturbative treatment of multiple scattering.
Consider a set of M scatterers, e.g. a collection of impurity atoms, that induce potential perturbations $\Delta V_1,\ \Delta V_2, \ldots,\ \Delta V_M$ with the total perturbation $\Delta V = \sum_i \Delta V_i$ producing a total $\mathcal{T}$-matrix denoted by $\mathcal{T}(E)$. Let $t_i(E)$ be the $\mathcal{T}$-matrix that the i-th scatterer would have if only the perturbation ΔV_i were present in the crystal. We seek an expansion formula of the total $\mathcal{T}$-matrix in terms of the individual t_i. The answer is (omitting the energy dependence)

$$\mathcal{T} = \sum_i t_i + \sum_{\substack{ij\\ i\neq j}} t_i \mathring{G}\, t_j + \sum_{\substack{ijk\\ i\neq j\\ j\neq k}} t_i \mathring{G}\, t_j \mathring{G}\, t_k + \cdots \tag{57}$$

which bears a multiple-scattering interpretation: the first term describes processes where the electron is scattered once by one of the scatterers and then leaves the scattering area; the second term describes scattering by successively two scatterers with an intermediate propagation determined by the crystal Hamiltonian; the third term describes three successive scattering events, etc. Note that the sequential scattering events are such that the electron is not scattered twice in a row by the same scatterer, i.e., no terms of the form $t_i \mathring{G} t_i$ appear since they are already contained in the t_i itself (although the electron can return to the scatterer i after intermediate scattering by k, i.e., the term $t_i \mathring{G} t_k \mathring{G} t_i$ is present). This is secured by the requirement $i \neq j$, $j \neq k$, etc., in the sums. This is in contrast to the expansion with respect to the potential, Eq. (56), where terms of successive interaction with the same scatterer are present. All these terms of the form $\Delta V_i \mathring{G} \Delta V_i$ of Eq. (56) are summed up in the t_i of Eq. (57).

Since in Eq. (57) there appear no successive terms form the same scatterer, it is convenient to express this by defining the off-site part of the Green function $\breve{G}_{ij}(E) = (1 - \delta_{ij})\mathring{G}(E)$. This is to be interpreted in the following sense. We assume that the potential perturbations ΔV_i are non-overlapping, e.g., situated at different atoms, and we also assume that we express the Green function in a localized, site-dependent basis set iL where i is the atomic site and L denotes the orbital. Then ΔV_i and t_i are matrices in L in the sub-block defining site i and otherwise zero (e.g. $(t_i)_{jL;j'L'} = (t_i)_{LL'}\delta_{ij}\delta_{jj'}$), $\mathring{G}$ is a full matrix, while $\breve{G}_{ij}(E)$ is the same as $\mathring{G}$ but with the $i = j$ (on-site) sub-blocks set to zero. We also define the sum of all individual t_i-matrices as the first term in Eq. (57), $t(E) = \sum_i t_i(E)$; $t(E)$ is thus a block-diagonal matrix. In this way, Eq. (57) can be brought to a closed form:

$$\mathcal{T} = t + t\breve{G}t + t\breve{G}t\breve{G}t + \cdots \tag{58}$$

$$= t + t\breve{G}\mathcal{T} \tag{59}$$

The size of the matrices in this equation is truncated by the number of perturbed sites M where $\Delta V_i \neq 0$ and by the truncation of the on-site local basis L. Eq. (59) bears analogy to Eq. (54) and can be solved by $\mathcal{T} = t\big[1 - \breve{G}\,t\,\big]^{-1}$. This point of view is the starting point of multiple scattering theory, including the Korringa-Kohn-Rostoker (KKR) method, that is beyond the scope of this manuscript.

When should one expect Eq. (57) or equivalently Eq. (58) to converge? If the single-site scattering matrix t is weak, then t is expected to be small and the series converges. However, even if t is large, the off-site Green function $\breve{G}$ is in many cases small enough that the series converges. Intuitively one can understand that $\mathring{G}(\boldsymbol{r}, \boldsymbol{r}'; E)$ will fall off with distance $r = |\boldsymbol{r} - \boldsymbol{r}'|$ in three dimensions: it describes a wave ψ originating at $\boldsymbol{r}'$ and propagating in all directions that must decay in amplitude as $\psi \sim r^{-1}$ to keep the probability current density $j \sim \psi^2 \sim r^{-2}$ integrated over a sphere of surface-area $4\pi r^2$ constant, independent of r. The same argument tells us that in two dimensions $\mathring{G}(\boldsymbol{r}, \boldsymbol{r}'; E)$ falls off slower, proportionally to $r^{-\frac{1}{2}}$, while in one dimension there is no falloff. Thus the expansion Eq. (58) is expected to converge rapidly in the case of scattering by a set of impurities in a crystal or less rapidly in the case of a thin film, especially if the impurities are not clustered together. Contrary to this, the expansion Eq. (56) is more restricted with respect to convergence because of the on-site multiple events $\Delta V_i(\mathring{G}\,\Delta V_i)^n$ whose convergence is subject to the weakness of ΔV_i. Finally, we note that the validity of the closed expressions Eq. (56) and Eq. (59) does not depend on the convergence properties of the corresponding series expansions.

3.4 Phase shifts

In analyzing the scattering properties of the system, it is convenient to choose a representation where the $\mathcal{S}$-matrix is diagonal. Since the $\mathcal{S}$-matrix is unitary, its eigenvalues will be in general complex with modulus of 1. Since it is energy conserving, its eigenvalues and eigenvectors will be characterized by the energy. Thus we may parametrize the $\mathcal{S}$-matrix eigenvalues as

$$s_\lambda(E) = \exp[2i\delta_\lambda(E)], \tag{60}$$

where $\delta_\lambda(E)$ are real parameters (so that $|s_\lambda(E)| = 1$) and λ is the set of quantum numbers other than the energy that are conserved during scattering and characterize the eigenvalue (e.g. in special cases λ could comprise the electron spin, orbital angular momentum or other symmetry properties). The δ_λ are called phase shifts for reasons that we will discuss below.
Bringing the $\mathcal{S}$-matrix in diagonal form means that we seek a superposition ψ_λ of in-states $|n\boldsymbol{k}\rangle \equiv |n\boldsymbol{k}_\parallel, E\rangle$ at energy E that is mapped onto itself (up to a constant) by the scattering process, i.e., we have

$$\psi_\lambda = \sum_{n\boldsymbol{k}} R_{\lambda;n\boldsymbol{k}}(E)\,|n\boldsymbol{k}_\parallel, E\rangle \rightarrow s_\lambda(E)\psi_\lambda = e^{2i\delta_\lambda(E)}\psi_\lambda \tag{61}$$

where $R_{\lambda;n\boldsymbol{k}}(E)$ is the transformation matrix making $\mathcal{S}$ diagonal, $R\mathcal{S}R^{-1} = \mathrm{diag}(s_\lambda)$. If Eq. (61) is satisfied, then the final state differs from the initial state only by a phase factor.
The term *phase shift* is justified when considering the case of a free-electron approximation scattered off a spherical potential. Then $|n\boldsymbol{k}\rangle = \Omega_{\mathrm{cryst}}^{-1/2}\exp(-i\boldsymbol{k}\cdot\boldsymbol{r})$, and neglecting spin-orbit coupling the quantum number λ corresponds to the orbital angular momentum number l (it is independent of the magnetic quantum number m or the spin). Then the eigenvectors of the $\mathcal{S}$-matrix are spherical waves whose radial part takes, asymptotically (for $r \rightarrow \infty$ where r is the distance from the scatterer), the form of free outgoing waves $\exp(ikr - \frac{1}{2}l\pi)$ in the absence of the potential. In the presence of the potential the waves are multiplied by the diagonal element s_l and take the form $\exp(2i\delta_l)\exp(ikr - \frac{1}{2}l\pi) = \exp[i(kr + 2\delta_l) - \frac{1}{2}l\pi]$, appearing as free waves but with their ampliude radially shifted shift by $2\delta_l/k$. From this aparent radial shift results the term phase shift for δ_l (see e.g. Ref. [2], Chap. 3.1.) In the case of free electrons the $\mathcal{S}$-matrix and the $\mathcal{T}$-matrix are diagonalized by the same transformation due to the high symmetry of the problem and the diagonal $\mathcal{T}$-matrix element becomes in the spherical-wave representation $t_l(E) = \pi k^{-1}[1 - s_l(E)]/2\pi i = k^{-1}[1 - \exp(2i\delta_l)]/2i = -(1/k)\sin\delta_l(E)\,\exp(i\delta_l(E))$ (the prefactor k^-1 comes from the transformation between the plane-wave and spherical-wave representation).

3.5 Relation of $\mathcal{T}$ to the density of states. Friedel sum rule

The scattering properties of a system are directly related to the density of states. In the physics of point defects in metals, this realization leads to the Friedel sum rule; and within multiple scattering theory it leads to approximations to total energy difference and interaction parameters by a perturbative expansion.
The starting point is the difference in the density of states between the perturbed and the unperturbed system $n(E)$ and $\mathring{n}(E)$ (see Eq. (21)):

$$\Delta n(E) := n(E) - \mathring{n}(E) = -\frac{1}{\pi}\mathrm{Im}\mathrm{Tr}[G(E) - \mathring{G}(E)]. \tag{62}$$

This expression can be further manipulated by observing that the definition of the Green function, Eq. (16), implies $G(E) = \partial(\ln G^{-1}(E))/\partial E$ (and analogously for $\mathring{G}(E)$) so that we obtain

$$\begin{aligned}\Delta n(E) &= -\frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\frac{\partial}{\partial E}[\ln G^{-1}(E) - \ln \mathring{G}^{-1}(E)] \\ &= -\frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\frac{\partial}{\partial E}\ln[\mathring{G}(E)G^{-1}(E)] \\ &= -\frac{1}{\pi}\mathrm{Im}\frac{\partial}{\partial E}\mathrm{Tr}\ln[1 - \mathring{G}(E)\,\Delta V]] \\ &= -\frac{1}{\pi}\mathrm{Im}\frac{\partial}{\partial E}\mathrm{Tr}\ln[\mathcal{T}^{-1}(E)\,\Delta V] \end{aligned} \quad (63)$$

In the first step we used the property $\mathrm{Tr}(\ln A + \ln B) = \mathrm{Tr}\ln(AB)$ for two matrices A and B that is valid because of the trace even if they do not commute (as long as the logarithms are defined).[6] The second step follows from the Dyson equation, Eq. (25), while the third step follows from Eq. (55). Integrating the latter expression, we obtain the difference in the integrated density of states,

$$\begin{aligned}\Delta N(E) = \int_{-\infty}^{E}\Delta n(E')dE' &= -\frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\ln[\mathcal{T}^{-1}(E)\,\Delta V]_{-\infty}^{E} \\ &= \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\ln\mathcal{T}(E)|_{-\infty}^{E} - \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\ln\,\Delta V|_{-\infty}^{E}. \end{aligned} \quad (64)$$

Since the potential is not energy dependent, the last term vanishes.[7] Also the term $\mathrm{Im}\mathrm{Tr}\ln\mathcal{T}(-\infty)$ gives no contribution: since the Hamiltonian has a lower bound in its spectrum,[8] $G(-\infty)$ vanishes $\mathcal{T}(-\infty) = \Delta V$ [see Eq. (52)]; and since ΔV is hermitian, $\mathrm{Tr}\ln\Delta V$ is real. Thus we arrive at the following relation of the $\mathcal{T}$-matrix to the difference in the integrated density of states between the reference and perturbed system:

$$\Delta N(E) = \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\ln\mathcal{T}(E) = \frac{1}{\pi}\mathrm{Im}\ln\det\mathcal{T}(E). \quad (65)$$

Differentiation gives for the difference in the density of states:[9]

$$\Delta n(E) = \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\left[\mathcal{T}^{-1}(E)\frac{d\mathcal{T}(E)}{dE}\right]. \quad (66)$$

We should note that Eq. (65) gives $\Delta N(E)$ only up to an arbitrary integer. The reason is that the logarithm is a multi-valued function in the complex plane. In calculations one has to follow the value of $\Delta N(E)$ over a dense enough mesh of energies starting from the band minimum in order to distinguish the correct integer part, because in a dense mesh a discontinuity of $\Delta N(E)$ by an integer would be clearly seen.

In the case of an impurity in a free-electron host crystal, the $\mathcal{T}$-matrix is l-dependent (see discussion in Sec. 3.4) and Eq. (65) reads

$$\Delta N(E) = \frac{2}{\pi}\sum_{lm}\ln T_l(E) = \frac{2}{\pi}\sum_{l}(2l+1)\delta_l(E) \quad (67)$$

[6] This can be shown by $\mathrm{Tr}(\ln A + \ln B) = \ln\det A + \ln\det B = \ln(\det A\det B) = \ln\det(AB) = \mathrm{Tr}\ln(AB)$, where $\det$ denotes the determinant.

[7] If the perturbation is energy-dependent, e.g., a self-energy, the term must be accounted for.

[8] Physically each atom has a finite number of bound states with energy $> -\infty$.

[9] Even if $\mathcal{T}$ does not commute with $d\mathcal{T}/dE$, the trace guarantees the validity of Eq. (66).

giving the displaced charge at energy E due to the impurity (a factor 2 accounting for the spin degeneracy and a factor has been introduced, since $\mathcal{T}$ is assumed to be spin independent in this case). At the Fermi energy, for reasons of charge neutrality, the displaced charge should match the difference of valence electrons between the impurity and host atoms $\Delta Z = Z_{\rm imp} - Z_{\rm host}$. From this follows the Friedel sum rule, $\Delta Z = \frac{2}{\pi}\sum_{lm} \ln T_l(E_{\rm F}) = \frac{2}{\pi}\sum_l (2l+1)\delta_l(E_{\rm F})$.
The formalism of the sum rule Eq. (65) allows us to calculate interaction parameters in a solid under the assumption that the change in total energy, under a perturbation, is to a good approximation equal to the change in the single-particle energies when the perturbation is weak enough. This assumption is known as *forcce theorem*. By the term *interaction parameters* we mean the following. In many cases certain excitations in a solid are described by a model Hamiltonian, e.g. magnetic excitations are described by a Heisenberg Hamiltonian $H = -\sum_{ij} J_{ij}\boldsymbol{M}_i \cdot \boldsymbol{M}_j$ where i and j run over all atoms and J_{ij} are the magnetic interaction parameters. In analogy one may assume a model with interaction parameters for inter-atomic forces, charge excitations, etc; or one may extend the model Hamiltonian to higher orders, e.g. considering bi-quadratic or 4-order terms among the moments in the magnetic excitations. We will discuss an application of this theory related to the Heisenberg model in Chapter C1 of the present volume.
In certain cases it happens that the $\mathcal{T}$-matrix, and thus value of the phase shift, change rapidly with energy. Then in a short energy interval around some $E_{\rm res}$ the phase shift increases by almost π. This occurs when the defect potential almost traps an electron due to a barrier, forming a *virtual bound state* or *resonance*. Then the Green function $G(E) = [z - H]^{-1}$ has a pole z_0 in the unphysical sheet, $\mathrm{Im}\sqrt{z_0} < 0$, with $\mathrm{Re}\, z_0 \approx E_{\rm res}$ in the continuous spectrum. This particular pole is close to the real axis causing the density of states, $-\mathrm{Im}G(E)/\pi$, to increase in the vicinity of $E_{\rm res}$, as can be seen by inspecting Eq. (66) and assuming that $d\mathcal{T}/dE$ is large. An example of resonance formation is the d state of a transition element embedded in a free-electron-like metal, e.g. Fe in Au. In the vicinity of $E_{\rm res}$ the scattering strength and the transition rate (defined below) increases, resulting e.g. in an increase of resistivity if $E_{\rm res}$ is at the Fermi energy.

3.6 Transition rate and optical theorem

In scattering calculations we are often interested in the transition rate $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$ from an initial state $|n\boldsymbol{k}\rangle$ to a final state $|n'\boldsymbol{k}'\rangle$. The concept of the transition rate is the following. Imagine that we prepare our system at some initial state $\psi_i(0)$ that propagates in time according to the time evolution $\psi_i(t) = e^{iHt}\psi_i(0)$ where H includes the perturbation ΔV that causes the scattering; usually $\psi_i(0)$ is a wavefunction propagating towards the perturbation region, but still far away from it. We seek the probability $P_{fi}(t) = |\langle\psi_f|\psi_i(t)\rangle|^2$ that a measurement finds the state ψ_f, where ψ_f is a given eigenfunction of the crystal Hamiltonian without the perturbation, $\mathring{H} = H - \Delta V$, measured far away from the scattering region. The transition rate is defined as the probability per unit time,

$$w_{fi} = \frac{dP_{fi}(t)}{dt}. \tag{68}$$

If the initial state is a wavepacket that is sharply peaked around a Bloch state $|n\boldsymbol{k}\rangle$ and the final state is a single Bloch wavefunction $|n'\boldsymbol{k}'\rangle$, then we obtain the transition rate $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$. In practice, knowledge of the scattering rate paves the way for electron-transport calculations e.g. by means of a semi-classical Boltzmann transport equation.
The scattering rate $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$ can be formally expressed in terms of the $\mathcal{T}$-matrix, $\mathcal{T}_{n'\boldsymbol{k}';n\boldsymbol{k}}$. To arrive at this expression we follow the derivation by Bohm [1] (Chap. XIV) but with simplifi-

cations because of the specific problem at hand (the derivation and the final formula by Bohm is valid for the general case). Starting from Eq. (68) we observe

$$w_{fi} = \langle\psi_f| \frac{d}{dt}\big(|\psi_i(t)\rangle\langle\psi_i(t)|\big)\,|\psi_f\rangle = -i\langle\psi_f|\,\big[H, |\psi_i(t)\rangle\langle\psi_i(t)|\big]\,|\psi_f\rangle \tag{69}$$

where we used the time-evolution law $dO(t)/dt = -i[H, O(t)]$ for the projection operator $O(t) = |\psi_i(t)\rangle\langle\psi_i(t)|$. Setting $H = \mathring{H} + \Delta V$ we obtain $w_{fi} = -i\langle\psi_f|\,\big[\,\Delta V, |\psi_i(t)\rangle\langle\psi_i(t)|\big]\,|\psi_f\rangle - i\langle\psi_f|\,\big[\mathring{H}, |\psi_i(t)\rangle\langle\psi_i(t)|\big]\,|\psi_f\rangle$. Since ψ_f is by assumption an eigenfunction of $\mathring{H}$, the Hamiltonian $\mathring{H}$ in the last term is replaced by the eigenenergy E_f of ψ_f and the commutator vanishes. For the remaining first term, $-i\langle\psi_f|\,\big[\,\Delta V, |\psi_i(t)\rangle\langle\psi_i(t)|\big]\,|\psi_f\rangle := -iC$, we use the fact that the commutator of two hermitian operators (here ΔV and $|\psi_i(t)\rangle\langle\psi_i(t)|$) is purely anti-hermitian and that the expectation value of an anti-hermitian operator $(\psi, C\psi)$ is purely imaginary.[10] Thus we find

$$w_{fi} = 2\mathrm{Im}\big[\langle\psi_f|\,\Delta V|\psi_i(t)\rangle\langle\psi_i(t)|\psi_f\rangle\big]. \tag{70}$$

Now we consider that $\psi_i(t)$ approaches asymptotically an incident wave $|n\boldsymbol{k}\rangle$ that evolves to $|n\boldsymbol{k};+\rangle$, i.e., we employ the form Eq. (40) for $\psi_i(t)$. In the limit that $\psi_i(t)$ becomes a monochromatic wave we can set $\psi_i(t) = |n\boldsymbol{k};+\rangle e^{-iE_{n\boldsymbol{k}}}$. We also accept that the final state is a Bloch eigenstate of $\mathring{H}$, $\psi_f = |n'\boldsymbol{k}'\rangle$. Then the terms on the rhs of Eq. (70) may be rewritten as follows:

$$\langle\psi_f|\,\Delta V|\psi_i(t)\rangle = \langle n'\boldsymbol{k}'|\,\Delta V|n\boldsymbol{k};+\rangle\, e^{-iE_{n\boldsymbol{k}}t} = \mathcal{T}_{n'\boldsymbol{k}';n\boldsymbol{k}}\, e^{-iE_{n\boldsymbol{k}}t} \tag{71}$$

by the definition of the $\mathcal{T}$-matrix (50) and

$$\begin{aligned}
\langle\psi_i(t)|\psi_f\rangle &= \langle n\boldsymbol{k};+|n'\boldsymbol{k}'\rangle e^{+iE_{n\boldsymbol{k}}t} \\
&= \big(\,\langle n\boldsymbol{k}|n'\boldsymbol{k}'\rangle + \langle n\boldsymbol{k};+|\;\Delta V[E_{n\boldsymbol{k}} - \mathring{H} - i0]^{-1}|n'\boldsymbol{k}'\rangle\,\big)e^{iE_{n\boldsymbol{k}}t} \\
&= \big(\,\Omega_{\rm rec}\, v_{n\boldsymbol{k}}\,\delta_{n'n}\,\delta(\boldsymbol{k}'_\parallel - \boldsymbol{k}_\parallel)\,\delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}}) \\
&\quad + \langle n\boldsymbol{k};+|\;\Delta V|n'\boldsymbol{k}'\rangle[E_{n\boldsymbol{k}} - E_{n'\boldsymbol{k}'} - i0]^{-1}\,\big)e^{iE_{n\boldsymbol{k}}t} \\
&= \big(\,\Omega_{\rm rec}\, v_{n\boldsymbol{k}}\,\delta_{n'n}\,\delta(\boldsymbol{k}'_\parallel - \boldsymbol{k}_\parallel)\,\delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}}) \\
&\quad + \mathcal{T}^*_{n\boldsymbol{k};n'\boldsymbol{k}'}\,[E_{n\boldsymbol{k}} - E_{n'\boldsymbol{k}'} - i0]^{-1}\,\big)e^{iE_{n\boldsymbol{k}}t}
\end{aligned} \tag{72}$$

where in the first step we used the Lippmann-Schwinger equation in the form (42), in the second step we used the orthonormality condition (38) and we exploited the fact that $|n'\boldsymbol{k}'\rangle$ is an eigenstate of $\mathring{H}$, and in the third step we used the definition of the $\mathcal{T}$-matrix (50). Inserting Eq. (71) and Eq. (72) into Eq. (70), and using the identity $\lim_{\eta\to 0}\mathrm{Im}(E - i\eta)^{-1} = \pi\delta(E)$, we arrive at the central result of this section[11]

$$w_{n'\boldsymbol{k}';n\boldsymbol{k}} = 2\pi\,|T_{n'\boldsymbol{k}';n\boldsymbol{k}}|^2\,\delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}}) \;+\; 2\mathrm{Im}\mathcal{T}_{n\boldsymbol{k};n\boldsymbol{k}}\,\Omega_{\rm rec}\,v_{n\boldsymbol{k}}\,\delta_{n'n}\delta(\boldsymbol{k}'_\parallel - \boldsymbol{k}_\parallel)\,\delta(E_{n'\boldsymbol{k}'} - E_{n\boldsymbol{k}}). \tag{73}$$

This result is one of the forms of Fermi's Golden Rule. From Eq. (73) it is immediately clear that the transition rate is energy-conserving. This is a result of the implicit assumption that the scattering target is merely a potential perturbation and not an object with internal degrees of freedom, as e.g. would be the case of scattering by phonons. In that case the total energy

[10] If A and B are hermitian then $C^\dagger := [A,B]^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A,B] = -C$, i.e., C is anti-hermitian. Then, for any state ψ, $(\psi, C\psi) = (C^\dagger\psi, \psi) = -(C\psi, \psi) = -(\psi, C\psi)^*$, i.e., $(\psi, C\psi)$ is purely imaginary.

[11] If one does not use atomic units with $\hbar = 1$, then the rhs of Eq. (73) must obtain a factor $1/\hbar$.

would be conserved, but the energy transition of the target would have to be accounted for. We also see that only the $\mathcal{T}$-matrix, and not the full $\mathcal{S}$-matrix, is sufficient to describe the scattering transition rate.
The result (73) is also given by Rodberg and Thaler (Ref. [2], Sec. 8.4.) who arrive at it by a different route, using time-dependent scattering theory. Following their argumentation, an additional result is derived. Since the set of all states $\{|n\boldsymbol{k}\rangle\}$ forms a complete basis, the probability for a transition to all of them together must be constant in time (and unity): the electron must be found in some state. This means that the sum of the transition rates over all possible final states, being the time derivative of the probability, must vanish. Thus we have, setting $\sum_{n'} \int d\boldsymbol{k}' w_{n'\boldsymbol{k}';n\boldsymbol{k}} = 0$ and using Eq. (73):

$$\mathrm{Im}\mathcal{T}_{n\boldsymbol{k};n\boldsymbol{k}} = -\frac{\pi}{\Omega_{\mathrm{rec}}} \sum_{n'} \int_{E_{n'\boldsymbol{k}'}=E_{n\boldsymbol{k}}} \frac{d\boldsymbol{k}'_{\parallel}}{v_{n'\boldsymbol{k}'_{\parallel}}} |T_{n'\boldsymbol{k}';n\boldsymbol{k}}|^2 \tag{74}$$

This formula is known as the *optical theorem*. It expresses conservation of probability during scattering and it can be derived independently from the unitarity of the $\mathcal{S}$-matrix (which also expresses conservation of probability). The relation $\mathcal{S}\mathcal{S}^\dagger = 1$ is written, using also the fact that the unit matrix is represented by $\delta_{n\boldsymbol{k};n'\boldsymbol{k}'} \to \Omega_{\mathrm{rec}}\delta(\boldsymbol{k}-\boldsymbol{k}')\delta_{nn'}$ and inserting Eq. (49),

$$\begin{aligned}
\Omega_{\mathrm{rec}}\,\delta(\boldsymbol{k}-\boldsymbol{k}')\delta_{nn'} &= \sum_m \int d\boldsymbol{q}\; \mathcal{S}_{n\boldsymbol{k};m\boldsymbol{q}}\mathcal{S}^*_{n'\boldsymbol{k}';m\boldsymbol{q}} \\
&= \sum_m \frac{1}{\Omega_{\mathrm{rec}}} \int \frac{d\boldsymbol{q}_{\parallel}}{v_{m\boldsymbol{q}}} \int dE \\
&\quad \times\delta(E-E_{n\boldsymbol{k}})\left[\Omega_{\mathrm{rec}}\, v_{n\boldsymbol{k}}\,\delta(\boldsymbol{q}_{\parallel}-\boldsymbol{k}_{\parallel})\delta_{mn} - 2\pi i\,\mathcal{T}_{n\boldsymbol{k};m\boldsymbol{q}}\right] \\
&\quad \times\delta(E-E_{n'\boldsymbol{k}'})\left[\Omega_{\mathrm{rec}}\, v_{n\boldsymbol{k}'}\,\delta(\boldsymbol{q}_{\parallel}-\boldsymbol{k}'_{\parallel})\delta_{mn'} + 2\pi i\,\mathcal{T}^*_{n'\boldsymbol{k}';m\boldsymbol{q}}\right] \\
&= \Omega_{\mathrm{rec}}\, v_{n\boldsymbol{k}}\,\delta(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}'})\,\delta(\boldsymbol{k}_{\parallel}-\boldsymbol{k}'_{\parallel})\,\delta_{nn'} \\
&\quad -2\pi i(\mathcal{T}_{n\boldsymbol{k};n'\boldsymbol{k}'}-\mathcal{T}^*_{n\boldsymbol{k};n'\boldsymbol{k}'})\,\delta(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}'}) \\
&\quad +4\pi^2 \sum_m \frac{1}{\Omega_{\mathrm{rec}}} \int \frac{d\boldsymbol{q}_{\parallel}}{v_{m\boldsymbol{q}}} \mathcal{T}_{n\boldsymbol{k};m\boldsymbol{q}}\mathcal{T}^*_{n'\boldsymbol{k}';m\boldsymbol{q}}\,\delta(E_{n\boldsymbol{k}}-E_{n'\boldsymbol{k}'})
\end{aligned} \tag{75}$$

The first term on the rhs cancels exactly the left-hand side (lhs) (applying Eq. (38)). The remaining terms are written as

$$\frac{1}{2i}(\mathcal{T}_{n\boldsymbol{k};n'\boldsymbol{k}'} - \mathcal{T}^*_{n\boldsymbol{k};n'\boldsymbol{k}'}) = \pi \sum_m \frac{1}{\Omega_{\mathrm{rec}}} \int \frac{d\boldsymbol{q}_{\parallel}}{v_{m\boldsymbol{q}}} \mathcal{T}_{n\boldsymbol{k};m\boldsymbol{q}}\mathcal{T}^*_{n'\boldsymbol{k}';m\boldsymbol{q}} \tag{76}$$

or in short

$$\frac{1}{2i}(\mathcal{T}-\mathcal{T}^\dagger) = \pi\,\mathcal{T}\mathcal{T}^\dagger, \tag{77}$$

valid for matrix elements between states on the same energy shell as implied by the energy-δ-function in the last term of Eq. (75). The expression (77) (or equivalently ((76)) is known as the *generalized optical theorem*. Setting $\boldsymbol{k} = \boldsymbol{k}'$ and $n = n'$ in Eq. (76), the optical theorem (74) follows.

4 Time reversal and space inversion symmetries

Among the possible symmetry operations obeyed by the crystal Hamiltonian, one may distinguish the special role of the time reversal and space inversion symmetries, as they play a role

in the degeneracy of the bands and the possibility for back-scattering. Additionally, due to the recent advances in the field of spin-orbit physics on surfaces [5, 6] and topologically protected surface states [7] there is increased interest in these two symmetries (and their braking).

4.1 General properties

The *space-inversion* operation $\mathcal{J}$ is the coordinate transformation $\boldsymbol{r} \to -\boldsymbol{r}$ around some fixed point in space. Thus, for any wavefunction $\psi(\boldsymbol{r})$,

$$\mathcal{J}\psi(\boldsymbol{r}) = \psi(-\boldsymbol{r}). \tag{78}$$

Obviously the space inversion operator is its own inverse, $\mathcal{J}^2 = 1$. It follows that the momentum operator transforms as $\hat{\boldsymbol{p}} = -i\boldsymbol{\nabla}_{\boldsymbol{r}} \to \mathcal{J}\hat{\boldsymbol{p}}\mathcal{J}^{-1} = -\hat{\boldsymbol{p}}$, while the kinetic energy operator $-\nabla^2$ remains unchanged. The orbital angular momentum $\hat{\boldsymbol{L}} = \boldsymbol{r} \times \hat{\boldsymbol{p}}$ and the spin $\frac{1}{2}\boldsymbol{\sigma}$, however, remain unchanged. If the crystal potential possesses space inversion symmetry around some point,[12] then the full Hamiltonian possesses space inversion symmetry. In particular, even in the presence of spin-orbit coupling[13]

$$\begin{aligned} H_{\rm soc} &= \frac{1}{4m^2c_0^2}\boldsymbol{\nabla}_{\boldsymbol{r}}V(\boldsymbol{r}) \times \hat{\boldsymbol{p}} \cdot \boldsymbol{\sigma} \\ &= \frac{1}{4m_r^2c_0^2}\frac{1}{r}\frac{dV(r)}{dr}\hat{\boldsymbol{L}} \cdot \boldsymbol{\sigma} \end{aligned} \tag{79}$$

the space-inversion symmetry is not lifted if $\mathcal{J}V\mathcal{J}^{-1} = V$, since $\mathcal{J}\hat{\boldsymbol{L}} \cdot \boldsymbol{\sigma}\mathcal{J}^{-1} = \hat{\boldsymbol{L}} \cdot \boldsymbol{\sigma}$. Assuming that the full crystal Hamiltonian $-\nabla^2 + V + H_{\rm soc}$ is invariant under space inversion, i.e., $[H, \mathcal{J}] = 0$, the Bloch eigenfunctions $\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = u_{n\boldsymbol{k}}(\boldsymbol{r})e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ (where $u_{n\boldsymbol{k}}$ is the periodic part of $\psi_{n\boldsymbol{k}}$) must be transformed to other Bloch eigenfunctions of H with the same energy: $H(\mathcal{J}\psi_{n\boldsymbol{k}}) = \mathcal{J}(H\psi_{n\boldsymbol{k}}) = E_{n\boldsymbol{k}}(\mathcal{J}\psi_{n\boldsymbol{k}})$. The transformed eigenfunction $\mathcal{J}\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = u_{n\boldsymbol{k}}(-\boldsymbol{r})e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}$ obviously belongs to the band-structure point $-\boldsymbol{k}$ (this is obvious from the phase $e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}$), so we may write $u_{n\boldsymbol{k}}(-\boldsymbol{r}) = u_{m,-\boldsymbol{k}}(\boldsymbol{r})$ for some band m at $-\boldsymbol{k}$ (there can be more than one bands at $-\boldsymbol{k}$ in case of a degeneracy; in particular we will examine the case of spin degeneracy below). Thus the space inversion maps in a sense $\boldsymbol{k}s \to -\boldsymbol{k}s$ and conserves the spin (or spin expectation value, if the spin is not a good quantum number due to spin-orbit coupling).
The *time reversal operator* $\mathcal{K}$ maps a wavefunction ψ to another in such a way that $\mathcal{K}\psi$ has the same density but opposite expectation value of the velocity, orbital angular momentum and spin. It must thus satisfy $(\mathcal{K}\,\psi, \hat{\boldsymbol{p}}\,\mathcal{K}\,\psi) = -(\psi, \hat{\boldsymbol{p}}\,\psi)$, $(\mathcal{K}\,\psi, \hat{\boldsymbol{L}}\,\mathcal{K}\,\psi) = -(\psi, \hat{\boldsymbol{L}}\,\psi)$, and $(\mathcal{K}\,\psi, \boldsymbol{\sigma}\,\mathcal{K}\,\psi) = -(\psi, \boldsymbol{\sigma}\,\psi)$, while it keeps the position unchanged, $(\mathcal{K}\,\psi, \hat{\boldsymbol{r}}\,\mathcal{K}\,\psi) = (\psi, \hat{\boldsymbol{r}}\,\psi)$. The term *time reversal* arises from the intuitive picture that a particle in the state $\mathcal{K}\psi$ would momentarily appear to be moving backward in time compared to a particle at state ψ. It can be shown that an operator with these properties cannot be unitary (see e.g. Ref. [1]). It is, however,

[12] This point should be either a nuclear position, as e.g. in monoatomic crystals or a position between two nuclei, as e.g. in the diamond structure.

[13] The spin-orbit coupling is inherent in the Dirac Hamiltonian and must be defined via the non-relativistic limit in the Schrödinger Hamiltonian. The first of Eq. (79) is the general form in the Schrödinger case while the second is an approximate form with r the distance from the closest atomic nucleus, taking into account that $H_{\rm soc}$ becomes non-negligible only very close to the nucleus. In the denominator, $m_r = m_{\rm e} + [E - V(r)]/2c_0^2$ indicates the r-dependent relativistic mass that partly cancels the unphysical divergence r^{-3} of $(1/r)dV/dr$ so that $H_{\rm soc} \sim r^{-1}$ at $r \to 0$.

anti-unitary, i.e., it conserves probability but is anti-linear: $\mathcal{K}(\alpha\psi_1 + \beta\psi_2) = \alpha^*\mathcal{K}\psi_1 + \beta^*\mathcal{K}\psi_2$, where α, β are complex numbers. A number of useful properties of the matrix elements of $\mathcal{K}$ is given by Yafet [4]. Here we delve into some of them that are needed below.

The demand that the velocity is reversed by $\mathcal{K}$ is met if we set $\mathcal{K}\,\psi = \psi^*$, i.e., $\mathcal{K}$ contains the complex conjugation that we call $\mathcal{K}_0$. This is easily seen by examining that the matrix element $(\psi^*, \hat{\boldsymbol{p}}\,\psi^*) = -(\psi, \hat{\boldsymbol{p}}\,\psi)$. Since $\hat{\boldsymbol{L}} = \boldsymbol{r} \times \hat{\boldsymbol{p}}$, reversing the velocity automatically reverses the orbital angular momentum. In order to reverse also the spin expectation value, $\mathcal{K}$ must include an interchange the up and down spin components. Demanding that this results in a reversal of the spin in all three directions, and including the complex conjugation it is found that the appropriate action of $\mathcal{K}$ is given by

$$\mathcal{K}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b^* \\ a^* \end{pmatrix} \tag{80}$$

This is satisfied if we set

$$\mathcal{K} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\mathcal{K}_0 \equiv -i\sigma_y\mathcal{K}_0 \tag{81}$$

The fact that $\mathcal{K}$ is anti-unitary means that it cannot be represented by a matrix (a matrix always represents a linear transformation and e.g. cannot perform the action of complex conjugation). Thus some commonly used manipulation rules do not apply, in particular the "action to a bra-vector on the left," i.e., we cannot write the adjoint of $\mathcal{K}\,|\psi\rangle$ as $\langle\psi|\,\mathcal{K}^\dagger$ and interpret $\mathcal{K}^\dagger$ by the usual transposition and complex-conjugation of $\mathcal{K}$.[14] The anti-unitarity also implies, for any two wavefunctions ψ_1 and ψ_2,

$$(\psi_1, \psi_2) = (\mathcal{K}\,\psi_1, \mathcal{K}\,\psi_2)^* = (\mathcal{K}\,\psi_2, \mathcal{K}\,\psi_1). \tag{82}$$

The $\mathcal{K}$ operator satisfies $\mathcal{K}^2 = -1$, from where it follows that $\mathcal{K}^{-1} = -\mathcal{K}$. By acting on a spinor function $\begin{pmatrix} a \\ b \end{pmatrix}$ it can be directly verified that $\mathcal{K}\,\hat{\boldsymbol{p}}\,\mathcal{K}^{-1} = -\hat{\boldsymbol{p}}$, from where it also follows that $\mathcal{K}\,\hat{\boldsymbol{L}}\,\mathcal{K}^{-1} = -\hat{\boldsymbol{L}}$. Additionally it can be straightforwardly verified that the spin operator transforms as $\mathcal{K}\,\boldsymbol{\sigma}\,\mathcal{K}^{-1} = -\boldsymbol{\sigma}$. The time-reversal operator commutes with the space-inversion operator, $[\mathcal{K}, \mathcal{J}] = 0$, as can be again verified directly by acting on a spinor function.

A Hamiltonian with no internal or external magnetic fields is invariant under the action of $\mathcal{K}$ even in the presence of spin-orbit coupling: the kinetic energy is invariant because $\mathcal{K}\,\hat{\boldsymbol{p}}^2\,\mathcal{K}^{-1} = \mathcal{K}\,\hat{\boldsymbol{p}}\,\mathcal{K}^{-1}\mathcal{K}\,\hat{\boldsymbol{p}}\,\mathcal{K}^{-1} = \hat{\boldsymbol{p}}^2$, the potential can be written in a purely spin-diagonal, real representation $V(\boldsymbol{r})$, and $H_{\rm soc}$ is also invariant because $\mathcal{K}\,\hat{\boldsymbol{L}}\cdot\boldsymbol{\sigma}\,\mathcal{K}^{-1} = \mathcal{K}\,\hat{\boldsymbol{L}}\,\mathcal{K}^{-1}\cdot\mathcal{K}\,\boldsymbol{\sigma}\,\mathcal{K}^{-1} = -\hat{\boldsymbol{L}}\cdot(-\boldsymbol{\sigma}) = \hat{\boldsymbol{L}}\cdot\boldsymbol{\sigma}$.

This invariance leads to a specific kind of degeneracy. If $\mathcal{K}\,H\mathcal{K}^{-1} = H$, and ψ is an eigenfunction of H with $H\,\psi = E\,\psi$, then $H\,(\mathcal{K}\,\psi) = \mathcal{K}\,\mathcal{K}^{-1}\,H\,\mathcal{K}\,\psi = \mathcal{K}\,H\,\psi = E\,(\mathcal{K}\,\psi)$, i.e., $\mathcal{K}\,\psi$ is also an eigenfunction of H, degenerate to ψ. Additionally, $\mathcal{K}\,\psi$ is also orthogonal to ψ, since $(\psi, \mathcal{K}\psi) = (\mathcal{K}^2\psi, \mathcal{K}\,\psi) = -(\psi, \mathcal{K}\,\psi)$, where we used Eq. (82) and $\mathcal{K}^2 = -1$. This means that in the presence of time-reversal invariance, the Hamiltonian eigenstates appear in orthogonal, degenerate pairs, ψ and $\mathcal{K}\,\psi$, named *Kramers pairs* and the two states are called *Kramers congugate* to one another; the degeneracy is named *Kramers degeneracy*.

[14] There is a meaningful definition of $\mathcal{K}^\dagger$ preserving the anti-liner property, but we avoid using it as it can lead to confusion in the manupulation of expressions.

For the crystal Bloch eigenfunctions, the Kramers degenerate of $\psi_{n\boldsymbol{k}} = u_{n\boldsymbol{k}}(\boldsymbol{r})e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ is $\mathcal{K}\,\psi_{n\boldsymbol{k}} = u^*_{n\boldsymbol{k}}(\boldsymbol{r})e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}$ that is again a Bloch function belonging to the band-structure point $-\boldsymbol{k}$ (by the phase factor $e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}$). It cannot, however, be the same as $\mathcal{J}\,\psi_{n\boldsymbol{k}}$, since $\mathcal{J}$ preserves the spin while $\mathcal{K}$ reverses it: $(\mathcal{J}\,\psi_{n\boldsymbol{k}}, \boldsymbol{\sigma}\,\mathcal{J}\,\psi_{n\boldsymbol{k}}) = (\psi_{n\boldsymbol{k}}, \boldsymbol{\sigma}\,\psi_{n\boldsymbol{k}}) = -(\mathcal{K}\,\psi_{n\boldsymbol{k}}, \boldsymbol{\sigma}\,\mathcal{K}\,\psi_{n\boldsymbol{k}})$. The two are, in fact, orthogonal to each other: $(\mathcal{J}\,\psi, \mathcal{K}\,\psi) = (\mathcal{K}\,\mathcal{J}\,\psi, \mathcal{K}^2\psi)^* = -(\mathcal{K}\,\mathcal{J}\,\psi, \psi)^* = -(\mathcal{K}\,\mathcal{J}^2\psi, \mathcal{J}\,\psi^*) = -(\mathcal{K}\,\psi, \mathcal{J}\,\psi)^* = -(\mathcal{J}\,\psi, \mathcal{K}\,\psi) = 0$, where we used Eq. (82), $\mathcal{K}^2 = 1$, $[\mathcal{K}, \mathcal{J}] = 0$, and $\mathcal{J}^2 = -1$. This means that, starting from a state $\psi_{n\boldsymbol{k}}$ and acting with $\mathcal{J}\mathcal{K}$ we obtain a Bloch eigenstate orthogonal to $\psi_{n\boldsymbol{k}}$ but at the same $\boldsymbol{k}$. This combined action of $\mathcal{C} := \mathcal{J}\mathcal{K}$ is called conjugation and obviously reverses the spin expectation value of $\psi_{n\boldsymbol{k}}$ due to the action of $\mathcal{K}$. In short,

$$H\,\mathcal{C}\,\psi_{n\boldsymbol{k}} := H\,\mathcal{J}\,\mathcal{K}\,\psi_{n\boldsymbol{k}} = E_{n\boldsymbol{k}}\,\mathcal{C}\,\psi_{n\boldsymbol{k}} \tag{83}$$

$$(\mathcal{C}\,\psi_{n\boldsymbol{k}}, \psi_{n\boldsymbol{k}}) = 0 \tag{84}$$

$$(\mathcal{C}\,\psi_{n\boldsymbol{k}}, \boldsymbol{\sigma}\,\mathcal{C}\,\psi_{n\boldsymbol{k}}) = -(\psi_{n\boldsymbol{k}}, \boldsymbol{\sigma}\,\psi_{n\boldsymbol{k}}) \tag{85}$$

An immediate consequence of the above is that in the presence of space-inversion and time-reversal symmetries, the band structure energies $E_{n\boldsymbol{k}}$ are at least twofold degenerate. This is reflected e.g. in the single-sheet Fermi surfaces of the alkali or noble metals. Conversely, if e.g. space inversion is not a symmetry of the system, the energy bands appear normally in non-degenerate, but close, pairs. This happens e.g. in the well-known cases of the conduction band of semiconductors in the zinc-blend structure, e.g. GaAs, or in the surface states of noble metals (where the presence of the surface breaks the inversion symmetry). It is worthwhile to note that this lifting of degeneracy by absence of space inversion happens because of spin-orbit coupling. Otherwise (and in the absence of magnetic fields) it is trivial to see that the two spins fully decouple and the same Hamiltonian is used for each spin separately, hence the exact same band structure occurs for each spin and the degeneracy is restored; in reality, however, the spin-orbit coupling is always present, even if it is very weak for light elements. To summarise, the absence of magnetic fields leads to a Kramers degeneracy between $\boldsymbol{k}$ and $-\boldsymbol{k}$ by time-reversal invariance, and the additional presence of space-inversion symmetry leads to a conjugation-degeneracy at $\boldsymbol{k}$ of two Bloch wavefunctions that have opposite spin expectation values.
An additional property of time-reversal symmetry that will be used below is the following [4]: Let $\hat{A}$ be an operator that transforms under time reversal to its hermitian conjugate. Then its matrix element between any two Kramers-conjugate states vanishes. I.e.,

$$\mathcal{K}\,\hat{A}\,\mathcal{K}^{-1} = \hat{A}^\dagger \;\Rightarrow\; (\psi, \hat{A}\,\mathcal{K}\,\psi) = 0. \tag{86}$$

To prove this we use Eq. (82) and $\mathcal{K}^{-1} = -\mathcal{K}$ and observe $(\psi, \hat{A}\,\mathcal{K}\,\psi) = (\mathcal{K}\,\hat{A}\,\mathcal{K}\,\psi, \mathcal{K}\,\psi) = -(\mathcal{K}\,\hat{A}\,\mathcal{K}^{-1}\,\psi, \mathcal{K}\,\psi) = -(\hat{A}^\dagger\psi, \mathcal{K}\,\psi) = -(\psi, \hat{A}\mathcal{K}\,\psi) = 0$.

4.2 Consequences of time-reversal and space-inversion in scattering

We now turn to certain consequences of the time-reversal and space-inversion symmetries and their braking in scattering properties. In particular, these symmetries affect the back-scattering amplitude, i.e. from $\boldsymbol{k}$ to $-\boldsymbol{k}$. We denote the Kramers conjugate of a Bloch state $|n\boldsymbol{k}\rangle$ by an overline, i.e., $|\overline{n\boldsymbol{k}}\rangle := \mathcal{K}\,|n\boldsymbol{k}\rangle$. Since (in the presence of time-reversal symmetry) this state belongs to the point $-\boldsymbol{k}$ of the band-structure, we denote the band that it belongs to also by an overline, so we write $|\overline{n}, -\boldsymbol{k}\rangle := |\overline{n\boldsymbol{k}}\rangle$. The adjoint of the state vector $|\overline{n\boldsymbol{k}}\rangle$ is symbolized by $\langle\overline{n\boldsymbol{k}}|$ or $\langle\overline{n}, -\boldsymbol{k}|$.

The first observation is that in the presence of time-reversal symmetry of both the crystal Hamiltonian $\mathring{H}$ and the impurity perturbation ΔV, i.e., if $\mathcal{K}\,\mathring{H}\,\mathcal{K}^{-1} = \mathring{H}\;\mathcal{K}\;\Delta V\,\mathcal{K}^{-1} = \Delta V$, and consequently $\mathcal{K}\,H\,\mathcal{K}^{-1} = H$, the scattering between Kramers conjugate states vanishes:

$$\langle n\boldsymbol{k}|\,\mathcal{T}\,|\overline{n},-\boldsymbol{k}\rangle \equiv \langle n,\boldsymbol{k}|\,\mathcal{T}\,|\overline{n\boldsymbol{k}}\rangle = 0. \tag{87}$$

This can be shown by a direct application of Theorem (86) and use of the expression (50) for the $\mathcal{T}$-matrix. From the form of the Green function, $G(E) = [E - H + i0]^{-1}$ it follows that $\mathcal{K}\,G(E)\,\mathcal{K}^{-1} = [E - \mathcal{K}H\mathcal{K}^{-1} - i0]^{-1} = [E - H + i0]^{-1} = G^{\dagger}(E)$ (since $H^{\dagger} = H$). Thus $\mathcal{K}\,\mathcal{T}\,\mathcal{K} = \mathcal{K}\,[\,\Delta V + \Delta V\,G(E)\;\Delta V]\,\mathcal{K}^{-1} = \Delta V + \Delta V\,\mathcal{K}\,G(E)\mathcal{K}^{-1}\,\Delta V = \Delta V + \Delta V\,G^{\dagger}(E)\;\Delta V = \mathcal{T}^{\dagger}$, since ΔV is hermitian. Applying Theorem (50) leads to Eq. (87).
From this we infer that, in a non-magnetic crystal, a non-magnetic defect in a non-magnetic crystal will not cause transitions between time-reverse states, while a magnetic defect will. A magnetic defect is characterized by the fact that ΔV contains a part with an exchange-correlation induced $\boldsymbol{B}$-field, i.e.,

$$\Delta V = \Delta V_0 + \boldsymbol{B}\cdot\boldsymbol{\sigma}. \tag{88}$$

But the $\boldsymbol{B}\cdot\boldsymbol{\sigma}$ part does not commute with the operator $\mathcal{K}$. Thus the proof of Eq. (87) does not hold in this case, and $\mathcal{T}_{n\boldsymbol{k},\overline{n}-\boldsymbol{k}} \neq 0$. In particular, using Eq. (81) and the relation $\mathcal{K}\,\boldsymbol{\sigma}\,\mathcal{K}^{-1} = -\boldsymbol{\sigma}$, we have $\mathcal{K}\,\boldsymbol{B}\cdot\boldsymbol{\sigma}\,\mathcal{K}^{-1} = -\boldsymbol{B}\cdot\boldsymbol{\sigma}$ (since $\boldsymbol{B}(\boldsymbol{r})$ is a real function of the coordinates only, it is not affected by the action of $\mathcal{K}$).
We may now see what is the effect of only a weak magnetic part of the impurity potential, i.e., expand the $\mathcal{T}$-matrix between Kramers-conjugate states to lowest order in $\boldsymbol{B}$. In doing this we must account for the effect of $\boldsymbol{B}$ on ΔV and on the impurity Green function $G(E)$ that enter the $\mathcal{T}$-matrix expression (50). For the Green function at $\boldsymbol{B} \neq 0$, $G_{\boldsymbol{B}}(E)$, we may use the Dyson equation with the system of the impurity with $\boldsymbol{B} = 0$ as a reference with Green function $G_0(E) = [E - \mathring{H} - \Delta V_0 + i0]^{-1}$. We have thus an expansion of the Green function:

$$G_{\boldsymbol{B}}(E) = G_0(E) + \boldsymbol{B}\cdot\boldsymbol{\sigma}\,G_0(E)\,\boldsymbol{B}\cdot\boldsymbol{\sigma} + \cdots, \tag{89}$$

where we employed Eq. (28) with $G_0(E)$ in the place of $\mathring{G}$ and $\boldsymbol{B}\cdot\boldsymbol{\sigma}$ in the place of ΔV. Using this expansion in the expression for the $\mathcal{T}$-matrix, Eq. (50), together with Eq. (88), we have

$$\begin{aligned}
\mathcal{T} &= \Delta V_0 + \boldsymbol{B}\cdot\boldsymbol{\sigma} + \big[\,\Delta V_0 + \boldsymbol{B}\cdot\boldsymbol{\sigma}\big]\big[G_0(E) + \boldsymbol{B}\cdot\boldsymbol{\sigma}\,G_0(E)\,\boldsymbol{B}\cdot\boldsymbol{\sigma} + \cdots\big]\big[\,\Delta V_0 + \boldsymbol{B}\cdot\boldsymbol{\sigma}\big] \\
&= \Delta V_0 + \Delta V_0\,G_0(E)\,\Delta V_0 + \boldsymbol{B}\cdot\boldsymbol{\sigma} + \Delta V_0\,G_0(E)\boldsymbol{B}\cdot\boldsymbol{\sigma} + \boldsymbol{B}\cdot\boldsymbol{\sigma}\,G_0(E)\,\Delta V_0 + O(B^2) \\
&\approx \mathcal{T}_0 + \boldsymbol{B}\cdot\sigma + \Delta V_0\,G_0(E)\boldsymbol{B}\cdot\boldsymbol{\sigma} + \boldsymbol{B}\cdot\boldsymbol{\sigma}\,G_0(E)\,\Delta V_0
\end{aligned} \tag{90}$$

at low $\boldsymbol{B}$, where $\mathcal{T}_0$ is the $\mathcal{T}$-matrix with only ΔV_0 as a perturbation. Thus we see that a weak $\boldsymbol{B}$-field enters linearly in the $\mathcal{T}$-matrix expression and quadratically in the transition rate, Eq. (73). For back-scattering between Kramers pairs the contribution by $\mathcal{T}_0$ vanishes, i.e., we obtain

$$\mathcal{T}_{n\boldsymbol{k};\,\overline{n},-\boldsymbol{k}} \approx \langle n\boldsymbol{k}|\,\big[\boldsymbol{B}\cdot\boldsymbol{\sigma} + \Delta V_0\,G_0(E)\boldsymbol{B}\cdot\boldsymbol{\sigma} + \boldsymbol{B}\cdot\boldsymbol{\sigma}\,G_0(E)\,\Delta V_0\big]\,|\overline{n},-\boldsymbol{k}\rangle \propto B \tag{91}$$

$$w_{n\boldsymbol{k};\,\overline{n},-\boldsymbol{k}} \propto B^2 \tag{92}$$

to lowest order in B.
A consequence of Eq. (87) (or Eq. (92)) is that in the absence of space-inversion symmetry in the crystal Hamiltonian there is no back-scattering from $\boldsymbol{k}$ to $-\boldsymbol{k}$ unless there are magnetic

impurities. This happens because without space-inversion symmetry there is no degeneracy in general at $\boldsymbol{k}$, except of course for certain high-symmetry lines in the Brillouin zone or points of accidental degeneracy that have, however, zero measure. The absence of back-scattering is indeed observed in systems of broken inversion symmetry and leads e.g. to absence of localization in the surface states of topological insulators [7]. There is, however, the possibility of back-scattering from $\boldsymbol{k}$ to $-\boldsymbol{k}+\delta\boldsymbol{k}$ for an arbitrarily small $\delta\boldsymbol{k}$ in the vicinity of $-\boldsymbol{k}$. This may be treated perturbatively with the $\boldsymbol{k}\cdot\boldsymbol{p}$ method.

The idea behind the $\boldsymbol{k}\cdot\boldsymbol{p}$ method is to find an approximation to the Bloch wavefunction at some $(n\boldsymbol{k})$ if the wavefunction at some close-by $(n\boldsymbol{k}_0)$ is known. Here we stress that both states must belong to the same band n for the method to work. The Schrödinger equation for the Bloch states $\psi_{n\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\cdot\boldsymbol{r}}u_{n\boldsymbol{k}}(\boldsymbol{r})$ can be rewritten with only the periodic part $u_{n\boldsymbol{k}}(\boldsymbol{r})$ as an unknown in the form [8]

$$\left[(\hat{\boldsymbol{p}}+\boldsymbol{k})^2 + V(\boldsymbol{r})\right] u_{n\boldsymbol{k}}(\boldsymbol{r}) = E_{n\boldsymbol{k}}\, u_{n\boldsymbol{k}(r)}. \tag{93}$$

Setting $\boldsymbol{k} = \boldsymbol{k}_0 + \delta\boldsymbol{k}$ and expanding $u_{n\boldsymbol{k}}$ in the basis set of periodic functions[15] $\{u_{n'\boldsymbol{k}_0}\}$ as $u_{n\boldsymbol{k}} = \sum_{n'} a_{n'} u_{n'\boldsymbol{k}_0}$, we obtain

$$\left[(\hat{\boldsymbol{p}}+\boldsymbol{k}_0+\delta\boldsymbol{k})^2 + V(\boldsymbol{r})\right] \sum_{n'} a_{n'} u_{n'\boldsymbol{k}_0}(\boldsymbol{r}) = E_{n\boldsymbol{k}} \sum_{n'} a_{n'} u_{n'\boldsymbol{k}_0}(\boldsymbol{r}). \tag{94}$$

Expanding to first order in $\delta\boldsymbol{k}$ we have $(\hat{\boldsymbol{p}}+\boldsymbol{k}_0+\delta\boldsymbol{k})^2 \approx (\hat{\boldsymbol{p}}+\boldsymbol{k}_0)^2 + 2\delta\boldsymbol{k}\cdot(\hat{\boldsymbol{p}}+\boldsymbol{k}_0)$ and projecting onto $u_{m\boldsymbol{k}_0}$ we get

$$\sum_{n'} a_{n'} \langle u_{m\boldsymbol{k}_0}| \left[(\hat{\boldsymbol{p}}+\boldsymbol{k}_0)^2 + 2\delta\boldsymbol{k}\cdot(\hat{\boldsymbol{p}}+\boldsymbol{k}_0) + V(\boldsymbol{r})\right] |u_{n'\boldsymbol{k}_0}\rangle = E_{n\boldsymbol{k}}\, a_m$$

which, realising that $\langle u_{m\boldsymbol{k}_0}| \left[(\hat{\boldsymbol{p}}+\boldsymbol{k}_0)^2 + V(\boldsymbol{r})\right] = E_{n'\boldsymbol{k}_0} \langle u_{m\boldsymbol{k}_0}|$, is rewritten as

$$\sum_{n'} a_{n'} \langle u_{m\boldsymbol{k}_0}| 2\delta\boldsymbol{k}\cdot(\hat{\boldsymbol{p}}+\boldsymbol{k}_0) + E_{m\boldsymbol{k}_0}\delta_{mn'} \big] |u_{n'\boldsymbol{k}_0}\rangle = E_{n\boldsymbol{k}}\, a_m.$$

Setting $\hat{\boldsymbol{p}}_{mn'} = \langle u_{m\boldsymbol{k}_0}|\hat{\boldsymbol{p}}|u_{n'\boldsymbol{k}_0}\rangle \equiv -i\langle u_{m\boldsymbol{k}_0}|\boldsymbol{\nabla}_{\boldsymbol{r}}|u_{n'\boldsymbol{k}_0}\rangle$ we arrive at

$$\sum_{n'} 2\delta\boldsymbol{k}\cdot\hat{\boldsymbol{p}}_{mn'} a_{n'} + E_{m\boldsymbol{k}_0} a_m = (E_{n\boldsymbol{k}} - 2\delta\boldsymbol{k}\cdot\boldsymbol{k}_0) a_m \tag{95}$$

as an algebraic eigenvalue equation for the coefficients $a_{n'}$ determining $u_{n\boldsymbol{k}}(\boldsymbol{r})$ and the eigenenrgy $E_{n\boldsymbol{k}}$ (remember that the $a_{n'}$ depend on n). As a cross-check we observe that for $\delta\boldsymbol{k} = 0$ we obtain $(E_{n\boldsymbol{k}_0} - E_{m\boldsymbol{k}_0})a_m = 0$ which is satisfied if $a_m = \delta_{mn}$ as expected.

Returning to the scattering problem, since we want to see how the $\mathcal{T}$-matrix $\mathcal{T}(E)$ changes in the vicinity of some $n\boldsymbol{k}_0$ we may set $E_{n\boldsymbol{k}} = E_{n\boldsymbol{k}_0} = E$ because we need to stay on the energy shell. Then Eq. (95) becomes

$$\sum_{n'} 2\delta\boldsymbol{k}\cdot\hat{\boldsymbol{p}}_{mn'} a_{n'} + E_{m\boldsymbol{k}_0} a_m = (E_{n\boldsymbol{k}_0} - 2\delta\boldsymbol{k}\cdot\boldsymbol{k}_0) a_m \tag{96}$$

Assuming that the $a_n u_{n\boldsymbol{k}_0}$ will dominate the expansion of $u_{n\boldsymbol{k}}$, which is reasonable since $\delta\boldsymbol{k}$ is assumed to be very small, we proceed by approximating the sum in Eq. (96) by only the n-th

[15] For any $\boldsymbol{k}$, the set $\{u_{n\boldsymbol{k}}(\boldsymbol{r})\}$, $n = 1, 2, \ldots$ forms a complete orthonormal system in the space of periodic functions in the primitive cell

term, i.e., just in the summation we set $a_{n'} = \delta_{nn'}$. Then we obtain $\delta\boldsymbol{k}\cdot\hat{\boldsymbol{p}}_{mn} + E_{m\boldsymbol{k}_0}a_m = (E_{n\boldsymbol{k}_0} - 2\delta\boldsymbol{k}\cdot\boldsymbol{k}_0)a_m$ which gives the approximate result

$$a_m \approx \begin{cases} \frac{\delta\boldsymbol{k}\cdot\hat{\boldsymbol{p}}_{mn}}{E_{n\boldsymbol{k}_0}-E_{m\boldsymbol{k}_0}}, & m \neq n \\ 1, & m = n \end{cases} \tag{97}$$

neglecting the term $2\delta\boldsymbol{k}\cdot\boldsymbol{k}_0$ in the denominator. The dependence on $n\boldsymbol{k}_0$ is implicit in the matrix elements $\hat{\boldsymbol{p}}_{mp}$ in addition to the denominator. The result of this whole $\boldsymbol{k}\cdot\hat{\boldsymbol{p}}$-maneuver is that we may write the Bloch wavefunction at $\boldsymbol{k}_0 + \delta\boldsymbol{k}$ as

$$|n, \boldsymbol{k}_0 + \delta\boldsymbol{k}\rangle \approx |n\boldsymbol{k}_0\rangle + \delta\boldsymbol{k}\cdot\sum_{m\neq n}\frac{\hat{\boldsymbol{p}}_{mn}}{E_{n\boldsymbol{k}_0} - E_{m\boldsymbol{k}_0}}|m\boldsymbol{k}_0\rangle \tag{98}$$

which is valid on the energy shell $E_{n\boldsymbol{k}_0} = E_{n,\boldsymbol{k}_0+\delta\boldsymbol{k}}$. Thus we see that the deviation of $|n, \boldsymbol{k}_0 + \delta\boldsymbol{k}\rangle$ from $|n, \boldsymbol{k}_0\rangle$ is linear in $\delta\boldsymbol{k}$.
Returning to the back-scattering $\mathcal{T}$-matrix element between Kramers conjugate pairs $|n\boldsymbol{k}\rangle$ and $|\overline{n}, -\boldsymbol{k}\rangle$, in the case of time-reversal symmetry, we arrive at the result that back-scattering in the immediate vicinity of $-\boldsymbol{k}$ is possible with the $\mathcal{T}$-matrix element behaving linear in $\delta\boldsymbol{k}$:

$$\begin{aligned} \mathcal{T}_{\overline{n},-\boldsymbol{k}+\delta\boldsymbol{k}\,;\,n\boldsymbol{k}} &\approx \delta\boldsymbol{k}\cdot\sum_{n'\neq\overline{n}}\frac{\hat{\boldsymbol{p}}_{nn'}}{E_{n\boldsymbol{k}} - E_{n'\boldsymbol{k}}}\langle n'\boldsymbol{k}|\big[\,\Delta V + \Delta V G(E_{n\boldsymbol{k}})\,\Delta V\big]|n\boldsymbol{k}\rangle \\ &\propto \delta k \end{aligned} \tag{99}$$

where the $\mathcal{T}$-matrix is explicitly written as $\Delta V + \Delta V G(E_{n\boldsymbol{k}})\,\Delta V$ in order to stress that these matrix elements are not any more between states of the same energy, i.e., they are not physical $\mathcal{T}$-matrix elements. The bottom line, however, is the linear behavior in $\delta\boldsymbol{k}$, which yields a quadratic increase of the transition rate around the forbidden Kramers-conjugate back-scattering point:

$$w_{\overline{n},-\boldsymbol{k}+\delta\boldsymbol{k}\,;\,n\boldsymbol{k}} \propto \delta k^2. \tag{100}$$

The expressions (99) and (100) are derived under the assumption that a single band $\overline{n}$ has the major contribution for small $\delta\boldsymbol{k}$ that led to the expansion coefficients (97) that breaks down close to degeneracy points.

5 Boltzmann transport equation

The transition rate w_{fi}, defined in Eq. (68) and given in Eq. (73), is one of the most useful quantities in scattering theory. In scattering experiments it is directly related to the scattering cross section, i.e., the scattered-particle flow measured by a detector at a certain direction normalized to the incident particle flow on the target. Since such experiments cannot be performed in the interior of a crystal, the cross section is not a directly measurable quantity. However, the transition rate paves the way to describe finite-lifetime effects and transport phenomena such as the electrical resistivity or the thermoelectric effect. A very fruitful theoretical approach for such phenomena is the semiclassical dynamics described by the Boltzmann equation. Within this approach, the crystal electrons are viewed as wavepackets that are sufficiently localized in space so that they form particles but also sufficiently localized in crystal-momentum to assign them with a single Bloch-state $|nk\rangle$ neglecting their dispersion. Under these conditions their

motion in the crystal follows the classical equations suggested by the Ehrenfest dynamics under external forces (electric and magnetic fields); during this motion, the internal forces (in particular the atomic potentials) are implicitly included in the band-structure via the group velocity and effective mass of the wavepackets. The motion is interrupted by collisions that re-destribute the electrons according to the transition rate $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$; after each collision the wavepacket is supposed to occupy a new state $|n'\boldsymbol{k}'\rangle$ and continue its motion under the external forces until the next collision, etc. In this way, one obtains a *distribution function* of the electrons on the Bloch states, $f(n\boldsymbol{k};T) = f_0(E_{n\boldsymbol{k}};T) + g(n\boldsymbol{k};T)$ where f_0 is the (equilibrium) Fermi function and $g(n\boldsymbol{k};T)$ is the deviation due to the external forces. Note that while f_0 depends on $|n\boldsymbol{k}\rangle$ only through the energy $E_{n\boldsymbol{k}}$, $f(n\boldsymbol{k};T)$ will normally depend on $|n\boldsymbol{k}\rangle$ explicitly.
The phase space of the semiclassical dynamics is the classical, in a sense, space of position and momentum, $\{\boldsymbol{r},\boldsymbol{p}\}$, with a minimal phase-space element compatible with the uncertainty relation for a wavepacket, $\Delta^3 r \Delta^3 p \sim (2\pi)^3$, which, however, is usually considered small enough so that continuum equations describe the electron trajectory in the phase space. (At some point the momentum is substituted by the crystal momentum.) The distribution function is therefore also $\boldsymbol{r}$ dependent with the implicit assumption that its spatial variation is significant at a length scale much larger than the wavepacket spread. The Boltzmann equation is a continuity equation in the phase space, describing the flow of the distribution function $f(\boldsymbol{r},\boldsymbol{p};T)$:

$$\frac{\partial f(\boldsymbol{r},\boldsymbol{p};T)}{\partial t} + \boldsymbol{\nabla}_{\boldsymbol{r}}(f\dot{\boldsymbol{r}}) + \boldsymbol{\nabla}_{\boldsymbol{p}}(f\dot{\boldsymbol{p}}) = \left.\frac{\partial f(\boldsymbol{r},\boldsymbol{p};T)}{\partial t}\right|_{\text{sc}}. \tag{101}$$

The lhs of the equation describes the flow of the distribution function in phase space, with the first term giving the rate of change at $(vr, n\boldsymbol{k})$, the second term giving the diffusion in $\boldsymbol{r}$ (containing the velocity $\dot{\boldsymbol{r}}$) and the third the diffusion in $\boldsymbol{k}$. In the absence of scattering the sum of the three terms is zero by the conservation of the number of electrons in the crystal. The right-hand side gives the rate of change of the distribution at $(vr, \boldsymbol{p})$ due to scattering processes. Expanding the divergence terms as

$$\boldsymbol{\nabla}_{\boldsymbol{r}}(f\dot{\boldsymbol{r}}) + \boldsymbol{\nabla}_{\boldsymbol{p}}(f\dot{\boldsymbol{p}}) = f(\boldsymbol{\nabla}_{\boldsymbol{r}}\dot{\boldsymbol{r}} + \boldsymbol{\nabla}_{\boldsymbol{p}}\dot{\boldsymbol{p}}) + \dot{\boldsymbol{r}}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}}f + \dot{\boldsymbol{p}}\cdot\boldsymbol{\nabla}_{\boldsymbol{p}}f$$

and using the Hamilton equations $\dot{\boldsymbol{r}} = \boldsymbol{\nabla}_{\boldsymbol{p}}H_c$ and $\dot{\boldsymbol{p}} = -\boldsymbol{\nabla}_{\boldsymbol{r}}H_c$ (where H_c is an assumed classical Hamiltonian, since we are still at the classical level), we see that the term $f(\boldsymbol{\nabla}_{\boldsymbol{r}}\dot{\boldsymbol{r}} + \boldsymbol{\nabla}_{\boldsymbol{p}}\dot{\boldsymbol{p}})$ vanishes and we arrive at the form

$$\frac{\partial f}{\partial t} + \dot{\boldsymbol{r}}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}}(f) + \dot{\boldsymbol{p}}\cdot\boldsymbol{\nabla}_{\boldsymbol{p}}(f) = \left.\frac{\partial f}{\partial t}\right|_{\text{sc}} \tag{102}$$

At this point the connection to the quantum mechanical description of the electrons is done. We substitute in Eq. (102) the velocity by the group velocity, $\dot{\boldsymbol{r}} \to \boldsymbol{v}_{n\boldsymbol{k}} = \boldsymbol{\nabla}_{\boldsymbol{k}}E_{n\boldsymbol{k}}$, and the momentum by the crystal momentum, $\boldsymbol{p} \to \boldsymbol{k}$ (using units with $\hbar = 1$). The semiclassical equation of motion of an electron in an external electric field $\boldsymbol{\mathcal{E}}$ and magnetic field $\boldsymbol{\mathcal{B}}$ is then given in terms of the Lorentz force

$$\dot{\boldsymbol{k}} = e(\boldsymbol{\mathcal{E}} + \frac{1}{c_0}\boldsymbol{v}_{n\boldsymbol{k}} \times \boldsymbol{\mathcal{B}}). \tag{103}$$

In addition, if we are interested in the description of a steady state, we set $\frac{\partial f}{\partial t} = 0$. Furthermore we allow for a temperature gradient in the crystal that is expressed through the term

$\boldsymbol{\nabla}_{\boldsymbol{r}} f(\boldsymbol{r}, \boldsymbol{k}; T)$ if we account for the dependence of f on T and the dependence of the Fermi energy or chemical potential μ on $\boldsymbol{r}$ via T or otherwise. Then the Fermi function becomes position dependent reading

$$f_0(\boldsymbol{r}, n\boldsymbol{k}; T) = \left[1 + \exp \frac{E_{n\boldsymbol{k}} - \mu(\boldsymbol{r})}{k_{\mathrm{B}} T(\boldsymbol{r})}\right]^{-1}. \tag{104}$$

Dropping terms of second or higher order in the electric field or in the temperature and chemical potential gradients,[16] we arrive at the linearized Boltzmann equation that has the form:[17]

$$\left(\frac{\partial f_0}{\partial E_{n\boldsymbol{k}}}\right) \boldsymbol{v}_{n\boldsymbol{k}} \cdot \left[e(\boldsymbol{\mathcal{E}} - \tfrac{1}{e}\boldsymbol{\nabla}_{\boldsymbol{r}}\mu) + \frac{E_{n\boldsymbol{k}} - \mu}{T}(-\boldsymbol{\nabla}_{\boldsymbol{r}} T)\right] + \frac{e}{c_0}(\boldsymbol{v}_{n\boldsymbol{k}} \times \boldsymbol{\mathcal{B}}) \cdot \boldsymbol{\nabla}_{\boldsymbol{k}} g = \left.\frac{\partial f}{\partial t}\right|_{\mathrm{sc}}. \tag{105}$$

The derivative $\partial f_0/\partial E$ arises by application of the chain rule when taking the gradients $\boldsymbol{\nabla}_{\boldsymbol{r}} f_0$ and $\boldsymbol{\nabla}_{\boldsymbol{k}}$. Since $\partial f_0/\partial E$ is a peaked function around the Fermi energy, in the limit of low temperatures it acts as a δ-function and pins the whole transport process in the vicinity of E_{F}.
The rhs of Eq. (105) is now treated by using the quantum-mechanical transition rate $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$ (73). The rate of change of the distribution at $(n\boldsymbol{k})$ is built up of two terms, one reducing the value of f for the electrons that are scattered out from this state to all others and one increasing the value of f for the electrons that are scattered from all other states into this one. Accounting for Fermi statistics, it is only possible to scatter an electron out of the state $|n\boldsymbol{k}\rangle$ if it is previously filled, therefore the transition rate $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$ is multiplied by the distribution function. We should also avoid the diagonal term $w_{n\boldsymbol{k};n\boldsymbol{k}}$ because it does not contribute to an exchange of electrons between different states; we denote this by a prime over the integral. We have:

$$\begin{aligned}
\left.\frac{\partial f}{\partial t}\right|_{\mathrm{sc}} &= \sum_{n'} \int' d\boldsymbol{k}' \left[f(\boldsymbol{r}, n'\boldsymbol{k}'; T)\, w_{n\boldsymbol{k};n'\boldsymbol{k}'} - f(\boldsymbol{r}, n\boldsymbol{k}; T)\, w_{n'\boldsymbol{k}';n\boldsymbol{k}}\right] \qquad (106)\\
&= \frac{2\pi}{\Omega_{\mathrm{rec}}} \sum_{n'} \int'_{E_{n'\boldsymbol{k}'} = E_{n\boldsymbol{k}}} \frac{d\boldsymbol{k}'_{\parallel}}{v_{n'\boldsymbol{k}'}} \Big[\, g(\boldsymbol{r}, n'\boldsymbol{k}'; T)\, |\mathcal{T}_{n\boldsymbol{k};n'\boldsymbol{k}'}|^2 - g(\boldsymbol{r}, n\boldsymbol{k}; T)\, |\mathcal{T}_{n'\boldsymbol{k}';n\boldsymbol{k}}|^2 \,\Big]
\end{aligned}$$

Here we set $f = f_0 + g$, we used the (energy-conserving) Eq. (73) for $(n\boldsymbol{k}) \neq (n'\boldsymbol{k}')$, and we took advantage of the property that f_0 depends on $(n\boldsymbol{k})$ only through the energy $E_{n\boldsymbol{k}}$ so that it can be pulled out of the constant-energy integral. Additionally we used the property $\sum_{n'} \int' d\boldsymbol{k}'\, w_{n\boldsymbol{k};n'\boldsymbol{k}'} = \sum_{n'} \int' d\boldsymbol{k}'\, w_{n'\boldsymbol{k}';n\boldsymbol{k}}$ to cancel the terms containing f_0. The latter property is justified by the following heuristic argument. The lhs of the equation gives the total rate of scattering into the state $|n\boldsymbol{k}\rangle$ from all other states, while the rhs gives the total rate of scattering out of the state $|n\boldsymbol{k}\rangle$ into all other states, independent of any occupation function f. If the two sides were different, then the weight of the state $|n\boldsymbol{k}\rangle$ should either increase or decrease with time even at equilibrium, which is unphysical.
It has been often suggested that there should be additional factors in Eq. (106) taking into account that an electron cannot be scattered into a state $|n\boldsymbol{k}\rangle$ if it is already filled. In this respect one would expect an expression of the form $\left[f_{n'\boldsymbol{k}'}(1 - f_{n\boldsymbol{k}})\, w_{n\boldsymbol{k};n'\boldsymbol{k}'} - f_{n\boldsymbol{k}}(1 - f_{n'\boldsymbol{k}'})\, w_{n'\boldsymbol{k}';n\boldsymbol{k}}\right]$. However, it has been shown by Kohn and Luttinger (Appendix of Ref. [9]) that this is not the case. If one drops terms of order g^2, the two expressions are identical in case that the system obeys the detailed balance condition, $w_{n\boldsymbol{k};n'\boldsymbol{k}'} = w_{n'\boldsymbol{k}';n\boldsymbol{k}}$, and the discrepancy does not show.

[16] We assume that terms of the form $g\boldsymbol{\mathcal{E}}$, $g\boldsymbol{\nabla}_{\boldsymbol{r}}\mu$, or $g\boldsymbol{\nabla}_{\boldsymbol{r}} T$ are of higher order and can be dropped.
[17] Note that the equation is not linearized with respect to the magnetic field.

However, the detailed balance condition does not necessarily hold in the presence of the spin-orbit interaction. Thus one should maintain the form without the $(1 - f_{n'\boldsymbol{k}'})$ and $(1 - f_{n\boldsymbol{k}})$ for calculations in systems with spin-orbit coupling.
The transition rate $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$ given by Eq. (73) corresponds to the case of a single defect. In a crystalline solid there is a concentration of defects that, within the Boltzmann-equation approach, are assumed to be randomly distributed. It is also assumed that the phase of the wavefunction is randomized after each scattering process, so that one may add the transition rates of all scattering processes instead of the transition amplitudes. Then the factor $w_{n'\boldsymbol{k}';n\boldsymbol{k}}$ should be scaled by the number of defects in the crystal (assuming that all defects are of the same type), $x_{\rm c}N_{\rm cr}$, where $x_{\rm c}$ is the defect concentration and $N_{\rm cr}$ the number of atoms (or primitive cells) in the crystal; i.e., we set $w_{n'\boldsymbol{k}';n\boldsymbol{k}} \to x_{\rm c}N_{\rm cr}\, w_{n'\boldsymbol{k}';n\boldsymbol{k}}$.
We can further manipulate Eq. (106) by defining the (concentration-normalized) relaxation time $\tau_{n\boldsymbol{k}}$ of state $|n\boldsymbol{k}\rangle$ by summing up the rate of transition to any other state via the expression[18]

$$x_{\rm c}N_{\rm cr}\,\frac{1}{\tau_{n\boldsymbol{k}}} = x_{\rm c}N_{\rm cr}\frac{2\pi}{\Omega_{\rm rec}}\sum_{n'}\int'_{E_{n'\boldsymbol{k}'}=E_{n\boldsymbol{k}}}\frac{d\boldsymbol{k}'_{\parallel}}{v_{n'\boldsymbol{k}'}}|\mathcal{T}_{n'\boldsymbol{k}';n\boldsymbol{k}}|^2 \tag{107}$$

By inserting Eq. (107) and Eq. (106) into Eq. (105), and accounting for the number of defects $x_{\rm c}N_{\rm cr}$, we obtain the following form of the Boltzmann equation:

$$\begin{aligned} g_{n\boldsymbol{k}}\, x_{\rm c}N_{\rm cr}\,\tau_{n\boldsymbol{k}}^{-1} \;=\; & -\left(\frac{\partial f_0}{\partial E_{n\boldsymbol{k}}}\right)\boldsymbol{v}_{n\boldsymbol{k}}\cdot\left[e(\boldsymbol{\mathcal{E}} - \tfrac{1}{e}\boldsymbol{\nabla}_{\boldsymbol{r}}\mu) + \frac{E_{n\boldsymbol{k}}-\mu}{T}(-\boldsymbol{\nabla}_{\boldsymbol{r}}T)\right] \\ & -\frac{e}{c_0}(\boldsymbol{v}_{n\boldsymbol{k}}\times\boldsymbol{\mathcal{B}})\cdot\boldsymbol{\nabla}_{\boldsymbol{k}}g_{n\boldsymbol{k}} + x_{\rm c}N_{\rm cr}\,\frac{2\pi}{\Omega_{\rm rec}}\sum_{n'}\int'_{E_{n'\boldsymbol{k}'}=E_{n\boldsymbol{k}}}\frac{d\boldsymbol{k}'_{\parallel}}{v_{n'\boldsymbol{k}'}}\,g_{n'\boldsymbol{k}'}\,|\mathcal{T}_{n\boldsymbol{k};n'\boldsymbol{k}'}|^2 \end{aligned} \tag{108}$$

with the obvious abbreviation $g_{n\boldsymbol{k}}$. Eq. (108) provides a self-consistent way of calculating the distribution function g. The lhs gives the rate of loss of electrons from the phase-space point $(\boldsymbol{r}, n\boldsymbol{k})$ due to scattering; the first and second term of the rhs give the rate of gain due to flow induced by the fields and temperature gradients; and the last term on the rhs gives the gain due to scattering in from all other states.
In order to arrive at a self-consistent solution of Eq. (108), an Ansatz is made on the form of g. We motivate this by an expansion of the distribution function f around the equilibrium f_0 in powers of the electric field:

$$f(\boldsymbol{r},\boldsymbol{k};T) = f_0(\boldsymbol{r},\boldsymbol{k};T) + \boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{\mathcal{E}}}f(\boldsymbol{r},\boldsymbol{k};T) + \cdots \tag{109}$$

The Ansatz is that to first order the electric field shifts the band energies, so that we may set $\boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{\mathcal{E}}}f = \frac{\partial f}{\partial E_{n\boldsymbol{k}}}\boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{\mathcal{E}}}E_{n\boldsymbol{k}} \approx \frac{\partial f_0}{\partial E_{n\boldsymbol{k}}}\boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{\mathcal{E}}}E_{n\boldsymbol{k}}$, dropping again the higher-order term $g\boldsymbol{\mathcal{E}}$. Thus we have

$$\begin{aligned} g(\boldsymbol{r},\boldsymbol{k};T) = f(\boldsymbol{r},\boldsymbol{k};T) - f_0(\boldsymbol{r},\boldsymbol{k};T) \;&=\; \frac{\partial f_0}{\partial E_{n\boldsymbol{k}}}\boldsymbol{\mathcal{E}}\cdot\nabla_{\boldsymbol{\mathcal{E}}}E_{n\boldsymbol{k}} \\ &=:\; -e\frac{\partial f_0}{\partial E_{n\boldsymbol{k}}}\boldsymbol{\mathcal{E}}\cdot\boldsymbol{\Lambda}_{n\boldsymbol{k}}. \end{aligned} \tag{110}$$

where we have introduced the *mean free path* $\boldsymbol{\Lambda}_{n\boldsymbol{k}}$. The result of the Ansatz is that we managed to write g in a linear form with respect to $\boldsymbol{\mathcal{E}}$ but also that we obtained a pre-factor $\frac{\partial f_0}{\partial E_{n\boldsymbol{k}}}$ that

[18] Eq. (107) contains a prefactor of $N_{\rm cr}/\Omega_{\rm rec}$ that diverges as $N_{\rm cr}^2$. This divergence is cancelled by the term $|\mathcal{T}_{n'\boldsymbol{k}';n\boldsymbol{k}}|^2$ because each $|n\boldsymbol{k}\rangle$ contains a normalization prefactor $1/\sqrt{N_{\rm cr}}$.

restricts g in a region of a few $k_{\mathrm{B}}T$ the vicinity of the Fermi level, which is physically expected since the deeper occupied levels are banned from contributing to the transport due to the Pauli exclusion principle, while the higher levels are unoccupied. The prefactor is also convenient for further manipulation of Eq. (108) where it already accompanies the term that does not depend on g.

The mean free path defined in Eq. (110) has also a physical interpretation as the average distance travelled between scattering events. During this "collisionless flight," the electron is accelerated by the electric field gaining energy $e\boldsymbol{\mathcal{E}} \cdot \boldsymbol{\Lambda}$, thus the electron energy changes from $E_{n\boldsymbol{k}}$ to $E_{n\boldsymbol{k}} - e\boldsymbol{\mathcal{E}} \cdot \boldsymbol{\Lambda}_{n\boldsymbol{k}}$.[19] This picture is consistent with the Ansatz leading to Eq. (110) and with the expression $\nabla_{\boldsymbol{\mathcal{E}}} E_{n\boldsymbol{k}} = -e\boldsymbol{\Lambda}_{n\boldsymbol{k}}$. Intuitively, the Fermi surface is displaced in the direction of $e\boldsymbol{\mathcal{E}}$ by an amount proportional to $e\boldsymbol{\mathcal{E}}$. The displacement does not increase indefinitely because scattering processes redistribute the electrons moving in the direction $e\boldsymbol{\mathcal{E}}$ to all other directions according to the stochastic process determined by the transition rate.

Substituting g from Eq. (110) into Eq. (108) we obtain an integro-differential equation for the mean free path. In the simpler case where the temperature and chemical-potential gradients vanish, the equation reads

$$\begin{aligned}(x_{\mathrm{c}}\,\boldsymbol{\Lambda}_{n\boldsymbol{k}} \cdot \hat{n}_{\boldsymbol{\mathcal{E}}})\, N_{\mathrm{cr}}\, \tau_{n\boldsymbol{k}}^{-1} \;&=\; (\boldsymbol{v}_{n\boldsymbol{k}} \cdot \hat{n}_{\boldsymbol{\mathcal{E}}}) - \frac{e}{c_0}\left(\boldsymbol{v}_{n\boldsymbol{k}} \times \frac{\boldsymbol{\mathcal{B}}}{x_{\mathrm{c}}}\right) \cdot \boldsymbol{\nabla}_{\boldsymbol{k}} (x_{\mathrm{c}}\,\boldsymbol{\Lambda}_{n\boldsymbol{k}} \cdot \hat{n}_{\boldsymbol{\mathcal{E}}}) \\ &\quad + \frac{2\pi N_{\mathrm{cr}}}{\Omega_{\mathrm{rec}}} \sum_{n'} \int'_{E_{n'\boldsymbol{k}'} = E_{n\boldsymbol{k}}} \frac{d\boldsymbol{k}'_{\parallel}}{v_{n'\boldsymbol{k}'}}\, |\mathcal{T}_{n\boldsymbol{k};n'\boldsymbol{k}'}|^2\, (x_{\mathrm{c}}\,\boldsymbol{\Lambda}_{n\boldsymbol{k}'} \cdot \hat{n}_{\boldsymbol{\mathcal{E}}}). \quad (111)\end{aligned}$$

The equation depends only on the direction of $\boldsymbol{\mathcal{E}}$, $\hat{n}_{\boldsymbol{\mathcal{E}}}$, and not on its magnitude, which is expected since we are only seeking the linear response of the system to the electric field. By choosing $\hat{n}_{\boldsymbol{\mathcal{E}}}$ along the three spatial directions we obtain three independent equations, one for each component of the vector $\boldsymbol{\Lambda}_{n\boldsymbol{k}}$. We also see that the impurity concentration can be included as a scaling constant of $\boldsymbol{\Lambda}$, as long as the magnetic field is also normalized to the concentration. This scaling is known as *Kohler's rule* [10] and has the following interpretation. A magnetic field causes the electron to travel on a continuous trajectory perpendicular to $\boldsymbol{\mathcal{B}}$ on the Fermi surface until it is interrupted by a scattering event. Higher fields cause a proportionally faster motion on the trajectory, while higher defect concentrations cause proportionally more frequent scattering events. Thus, scaling up the defect concentration has the same effect as scaling down the magnetic field.

In the absence of a magnetic field, Eq. (111) can be fairly easily solved with iterations, while for $\boldsymbol{\mathcal{B}} \neq 0$ the solution is more complicated because of the derivative term $\boldsymbol{\nabla}_{\boldsymbol{k}}\boldsymbol{\Lambda}$. We should further note that the mean free path is a function of the magnetic field, $\boldsymbol{\Lambda}_{n\boldsymbol{k}}(\boldsymbol{\mathcal{B}})$, and Eq. (111) must be solved explicitly for any value of $\boldsymbol{\mathcal{B}}$. However, at low fields an expansion of $\boldsymbol{\Lambda}_{n\boldsymbol{k}}(\boldsymbol{\mathcal{B}})$ in powers of $|\boldsymbol{\mathcal{B}}|$ is possible, leading to a hierarchy of equations that give the expansion parameters [11].

Once the mean free path has been found, the current density can be calculated by means of the

[19] The negative sign arises because a constant electric field $\boldsymbol{\mathcal{E}}$ acting on a charge e corresponds to a potential $-e\boldsymbol{\mathcal{E}} \cdot \boldsymbol{r}$. Substituting $\boldsymbol{\Lambda}$ for $\boldsymbol{r}$ we arrive at Eq. (110). In other words, the states with group velocity parallel to $e\boldsymbol{\mathcal{E}}$ should decrease their energy and increase their occupation.

distribution function:

$$\begin{aligned} \boldsymbol{j}(\boldsymbol{r}) &= \frac{e}{\Omega_{\rm cryst}} \sum_n \int d\boldsymbol{k}\, \boldsymbol{v}_{n\boldsymbol{k}}\, f(\boldsymbol{r}, n\boldsymbol{k}; T) \\ &= \frac{e}{\Omega_{\rm cryst}\Omega_{\rm rec}} \sum_n \int dE \int_{E_{n\boldsymbol{k}}=E} \frac{d\boldsymbol{k}_\parallel}{v_{n\boldsymbol{k}}}\, \boldsymbol{v}_{n\boldsymbol{k}}\, g(\boldsymbol{r}, n\boldsymbol{k}; T) \\ &= x_{\rm c} \frac{e^2}{8\pi^3} \sum_n \int dE \left(-\frac{\partial f_0}{\partial E}\right) \int_{E_{n\boldsymbol{k}}=E} \frac{d\boldsymbol{k}_\parallel}{v_{n\boldsymbol{k}}} \boldsymbol{v}_{n\boldsymbol{k}} (\boldsymbol{\Lambda}_{n\boldsymbol{k}} \cdot \boldsymbol{\mathcal{E}}) \end{aligned} \tag{112}$$

with $\Omega_{\rm cryst} = N_{\rm cr} 8\pi^3/\Omega_{\rm BZ}$ the crystal volume. In the first step we substituted $f = f_0 + g$ and we used the fact that the equilibrium distribution does not contribute to the current. From here, the conductivity tensor is readily obtained as the prefactor to the electric field:

$$\sigma_{ij} = x_{\rm c} \frac{e^2}{8\pi^3} \sum_n \int dE \left(-\frac{\partial f_0}{\partial E}\right) \int_{E_{n\boldsymbol{k}}=E} \frac{d\boldsymbol{k}_\parallel}{v_{n\boldsymbol{k}}} (\boldsymbol{v}_{n\boldsymbol{k}})_i\, (\boldsymbol{\Lambda}_{n\boldsymbol{k}})_j \tag{113}$$

with $i, j = x, y, z$. At low temperatures, $-\partial f_0/\partial E \to \delta(E - E_{\rm F})$, confining the transport properties to the Fermi energy. From the diagonal elements of the tensor σ_{ij} one obtains the resistivity and magneto-resistance, while from the off-diagonal elements one obtains the Hall conductivity.

Acknowledgments

I would like to thank Philipp Rüßmann for a critical reading of the manuscript.

Appendix

Reduction of the wavepacket to the free wavepacket at $t \to -\infty$

Here we show that the second term of the Lippmann-Schwinger equation for a wavepacket vanishes at very early times [see the discussion following Eq. (41)]. The term that has to vanish is $\text{(II)} = \lim_{\eta\to 0} \int d\boldsymbol{k} A(\boldsymbol{k}-\boldsymbol{k}')[E_{n\boldsymbol{k}} - H + i\eta]^{-1}\, \Delta V |n\boldsymbol{k}\rangle e^{-iEt}$, so that the wavepacket Eq. (40) reduces to the free wavefunction Eq. (39) at $t \to -\infty$.
We follow arguments given by Bohm (Ref. [1], Chap. XV). We first make a transformation from $\boldsymbol{k} = (\boldsymbol{k}_\parallel, \boldsymbol{k}_\perp)$ to $(\boldsymbol{k}_\parallel, E)$ to obtain the energy as an integration variable. We also assume that the wavepacket amplitude A can be separated in a $\boldsymbol{k}_\parallel$-part $A_\parallel(\boldsymbol{k}_\parallel - \boldsymbol{k}'_\parallel)$ and in a $\boldsymbol{k}_\perp$-part $A_\perp(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp) = A_{\rm E}(E - E')$, where E is the energy corresponding to $\boldsymbol{k}_\perp$ and E' the energy corresponding to $\boldsymbol{k}'_\perp$. The subscript E in $A_{\rm E}$ is used to distinguish it from the function A. Making the transformation to the $(\boldsymbol{k}_\parallel, E)$ variables we obtain

$$\text{(II)} = \lim_{\eta\to 0} \int d\boldsymbol{k}_\parallel A_\perp(\boldsymbol{k}_\perp - \boldsymbol{k}'_\perp) \int dE A_{\rm E}(E - E')[E - H + i\eta]^{-1}\, \Delta V |n\boldsymbol{k}_\parallel, E\rangle e^{-iEt} \tag{114}$$

where we used the renotation $|n\boldsymbol{k}_\parallel, E\rangle = |n(\boldsymbol{k}_\parallel, \boldsymbol{k}_\perp)\rangle =: |n\boldsymbol{k}\rangle$. We examine the time evolution of a wavepacket that is formed by the states $|n\boldsymbol{k}; +\rangle$, use the Lippmann-Schwinger equation to

substitute $|n\boldsymbol{k};+\rangle$, and see that there remains a free-wave part and a scattered wave part that vanishes for $t \to -\infty$. Basically this procedure leads from the energy-dependent Lippmann-Schwinger equation, Eq. (32), to the time-dependent Lippmann-Schwinger equation, Eq. (13). In a brief outline, consider a wavepacket of the form (40)

$$\begin{aligned} \psi(t) &= \int d\boldsymbol{k}\, A(\boldsymbol{k}-\boldsymbol{k}')\, |n\boldsymbol{k};+\rangle\, e^{-iE_{n\boldsymbol{k}}t} \\ &= \int d\boldsymbol{k}_\parallel\, A_\parallel(\boldsymbol{k}_\parallel - \boldsymbol{k}'_\parallel) \int dE\, A_{\rm E}(E-E')\, |n\boldsymbol{k}_\parallel E;+\rangle\, e^{-iEt} \end{aligned} \tag{115}$$

where the conventions of the previous paragraph were used in the notation, i.e., $|n\boldsymbol{k}_\parallel E;+\rangle = |n\boldsymbol{k}_\parallel \boldsymbol{k}_\perp;+\rangle = |n\boldsymbol{k};+\rangle$ with $E = E_{n\boldsymbol{k}}$. By virtue of the Lippmann-Schwinger equation (41) the above expression becomes

$$\begin{aligned} \psi(t) &= \int d\boldsymbol{k}_\parallel\, A_\parallel(\boldsymbol{k}_\parallel - \boldsymbol{k}'_\parallel) \int dE\, A_{\rm E}(E-E')\, |n\boldsymbol{k}_\parallel E\rangle\, e^{-iEt} \\ &\quad + \int d\boldsymbol{k}_\parallel\, A_\parallel(\boldsymbol{k}_\parallel - \boldsymbol{k}'_\parallel) \int dE\, A_{\rm E}(E-E')\, [E-H+i\eta]^{-1}\, \Delta V |n\boldsymbol{k}_\parallel E\rangle\, e^{-iEt}. \end{aligned} \tag{116}$$

The first term on the rhs expresses just the time-evolution of the free wave-packet $\mathring{\psi}(t)$ and is identical to Eq. (39). The second part of the rhs contains a Fourier transform from the energy domain to the time domain of a product of two functions, $f(E) = [E-H+i\eta]^{-1}\,\Delta V$ and $g(E) = A_{\rm E}(E-E')|n\boldsymbol{k}_\parallel E\rangle$. Each of them has its individual Fourier transform, in particular[20] $\tilde{f}(t) = \int dE[E-H+i\eta]^{-1}\,\Delta V\, e^{-iEt} = -i\theta(t)e^{-iHt}$ and $\tilde{g}(t) = \int dE\, f(E)e^{-iEt} = A_{\rm E}(E-E')\, |n\boldsymbol{k}_\parallel E\rangle\, e^{-iEt}$ which is basically just the free wavepacket $\mathring{\psi}(t)$ (implying the $d\boldsymbol{k}_\parallel$-integration). But the Fourier transform of a product of functions f and g is the convolution of the individual Fourier transforms: $\int f(E)g(E)e^{-iEt} = \int dt' \tilde{f}(t-t')\tilde{g}(t') = \int dt' i\theta(t-t')e^{-iH(t-t')}\mathring{\psi}(t')$ which is just the second part of the rhs in Eq. (13).
At very early times ($t \to -\infty$), the latter term, $\int dt' i\theta(t-t')e^{-iH(t-t')}\mathring{\psi}(t')$, vanishes because of the θ step function, under the condition that the product $\Delta V\mathring{\psi}(t')$ also vanishes for $t' \to -\infty$. This is the case if the potential perturbation ΔV is localized in space and the wavepacket $\mathring{\psi}$ is prepared far away from the region of ΔV as is usually the case in a scattering situation, or if ΔV is switched on slowly with time so that it vanishes in the remote past, irrespective of its range; otherwise the wavepacket is never far enough from the perturbation to be considered as a free incoming wave. It is worthwhile to note that the Coulomb potential, falling off as $1/r$, is not localized enough to meet these conditions, therefore the description scattering off charged defects in solids is a challenge. However, at the end there is always some degree of screening of the charge by mobile carriers in metals or semiconductors or by the dielectric function in insulators allowing for the use of standard theory.

[20]The Fourier transform of $\lim_{\eta\to 0}(E-E'\pm i\eta)^{-1}$ with respect to E is $\mp ie^{iE't}\theta(\pm t)$.

References

[1] Arno Bohm, *Quantum Mechanics: Foundations and Applications*, Third Edition, Springer (2001).

[2] Leonard S. Rodberg and R. M. Thaler, *Introduction to the Quantum Theory of Scattering*, Academic Press (1967).

[3] P. Mavropoulos and N. Papanikolaou, *The Green function method of Korringa, Kohn and Rostoker. I. Electronic structure of periodic systems*, in *Computational Nanoscience: Do it yourself!*, edited by J. Grotendorst, S. Blügel, and D. Marx, NIC series vol. 31, Forschungszentrum Jülich (2006); http://webarchiv.fz-juelich.de/nic-series/volume31/volume31.html

[4] Y. Yafet, *g factors and spin-lattice relaxation of conduction electrons*, in Solid State Physics, edited by F. Seitz and D. Turnbull, vol. 14, 1 (1963).

[5] S. LaShell, B. A. McDougall, and E. Jensen, *Spin Splitting of an Au(111) Surface State Band Observed with Angle Resolved Photoelectron Spectroscopy*, Phys. Rev. Lett. **77**, 3419 (1996).

[6] J. Henk, M. Hoesch, J. Osterwalder, A. Ernst, and P. Bruno, *Spinorbit coupling in the L-gap surface states of Au(111): spin-resolved photoemission experiments and first-principles calculations*, J. Phys.: Condens. Matter **16**, 7581 (2004).

[7] M. Z. Hasan and C. L. Kane, *Colloquium: Topological insulators*, Rev. Mod. Phys. **82**, 3045 (2010).

[8] Neil W. Ashcroft and N. David Mermin, *Solid State Physics*, Cengage 1976.

[9] W. Kohn and J. M. Luttinger, *Quantum Theory of Electrical Transport Phenomena*, Phys. Rev. **108**, 590 (1957).

[10] M. Kohler, *The magnetic resistance changes of pure metals*, Ann. Phys. **32**, 211 (1938).

[11] P. Mavropoulos and N. Stefanou, *Low-field galvanomagnetic properties of aluminium-based dilute alloys*, J. Phys: Condens. Matter **9**, 8997 (1997).

A 6 Geometric Phases and Topological effects[1]

Yuriy Mokrousov and Frank Freimuth
Peter Grünberg Institut (PGI-1) and Institute for Advanced Simulation (IAS-1)
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Since the beginning of the eighties, and especially since the seminal work by Berry [1], the mathematical concepts of geometry, topology, geometric phase and topological characterization of so-called fibre bundles, rapidly entered various aspects of condensed matter physics such as, among others, electric polarization, Hall effect in insulators and metals, transport properties at surfaces and magnetization dynamics. Especially recently, triggered by discovery of topological and spontaneous Chern insulators, the topological classification of solids based on quantities derived from their geometrical properties has become a very common tool in characterization of physical properties of metals and insulators. Moreover, interplay of complex magnetism with topological properties of solids is being studied very intensively nowadays.

It is impossible to review all of the issues listed above with a descent degree of depth within the format of comparatively short lecture notes. We have thus decided to focus on selected aspects of the geometry-related topics in condensed matter physics, trying to keep our manuscript as concise and as self-contained as possible. An interested reader should be able to follow our notes from the beginning to the end with minimal reference to other sources. We start by discussing the mathematical foundations of the Berry, or, geometric phase, formulate the adiabatic approximation for quantum dynamics and introduce fundamental concepts such as Berry connection, Berry curvature, gauge freedom, parallel transport and first Chern number. The physical examples we choose to apply the introduced concepts to are the Aharonov-Bohn effect and spin-$\frac{1}{2}$ in magnetic field, which prove to be of great importance for understanding the material in the rest of the notes. We further show in detail the geometrical and topological nature of electric polarization in insulators, bringing to attention its relation to the Chern number. In a separate chapter we discuss, referring to simple arguments, the emergence of the Chern insulators in two-dimensional reciprocal space, and the interplay between various versions of, quantized and not, Hall effects taking place in metals and insualtors. We derive the expression for the velocity of a state due to time-evolution of the quantum system and express it in geometrical terms, and use it to arrive at the equations of motion which govern the dynamics of electrons in a solid in response to electro-magnetic fields and general perturbations. We discuss the consequences of these equations for such properties as Hall conductance and orbital magnetization. Finally, as an example of a system with properties dependent on a parameter slowly varying in real space we choose a solid with spatially varying magnetization direction. We show how reformulation of the problem in geometrical terms can be used to explain the rise of emergent magnetic field in such systems, and the emergence of the topological Hall effect. We conclude with the purely geometric semiclassical derivation of the Dzyaloshinskii-Moriya interaction.

There is a number of books and reviews in which most of the aspects presented here can be found, and in these sources an interested reader can find more details and more profound discussions. Namely, we refer to the books by Bohm [2] and Nakahara [3] for a mathematically rigorous discussion of geometric phase, fibre bundle theory and dynamics of quantum systems, briefly outlined in section 2. The subject of electric polarization presented in section 3.2 has been discussed in depth in [4]. Various aspects of topological and Chern insulators are concisely and transparently presented and discussed in [5]. The best review of the issues of Berry phase related electron dynamics in external perturbations can be found in Ref. [6]. A good introduction to the topological Hall effect is given in [7]. Overall, we are also grateful to Hongbin Zhang, Robert Bamler, Achim Rosch, Gustav Bihlmayer and Stefan Blügel for multiple discussions on the subject, which helped in streaming the manuscript and making many things more transparent.

Finally, we remark that, for reasons of simplification, in the following we assume that $m_e = \hbar = e = c = 1$. Most of the expressions which include the prefactors and which should be used for practical evaluation of the quantities, can be found in the sources listed above.

2 General Theory

2.1 Berry phase and adiabatic evolution

The origins of the Berry phase lie in the dynamics of a quantum system described by a Hamiltonian $H(\lambda)$ which bares a parametric dependence on a parameter λ. It is normally assumed that λ lies on a certain differentiable manifold M, $\lambda \in M$. We assume that the Hamitonian $H(\lambda)$ as well as its descrete eigenspectrum $\{\varepsilon_n(\lambda)\}$ are smooth and unique functions of λ everywhere on M. For each λ we denote by $\{|n\lambda\rangle \in \mathcal{H}\}$ a set of "instantaneous" solutions of

$$H(\lambda)\,|n\lambda\rangle = \varepsilon_n(\lambda)\,|n\lambda\rangle\,. \tag{1}$$

It is important to realize that, in contrast to Hamiltonian and eigenvalues, each function $|n\lambda\rangle : M \to \mathcal{H}$ can be chosen to be smooth only over certain parts of M (called *patches*), but not necessarily over the whole of M. Imagine two patches O_1 and O_2 in M, $O_1 \cap O_2 \neq \varnothing$, with two sets of smooth functions $\{|n\lambda\rangle\}$ and $\{|n\lambda\rangle'\}$ on them. Then for any $\lambda \in O_1 \cap O_2$ we know that $|n\lambda\rangle$ and $|n\lambda\rangle'$ can differ only by a complex phase which is the element of group $\mathbb{U}(1)$:

$$|n\lambda\rangle' = e^{i\zeta_n(\lambda)}\,|n\lambda\rangle\,. \tag{2}$$

Also on the patches themselves we can always switch from $\{|n\lambda\rangle\}$ to the "alternative" functions $\{|n\lambda\rangle'\}$ via the *gauge transformation* given by (2) with arbitrary functions $\{\zeta_n(\lambda)\}$ being smooth on corresponding patches. Indeed, both $\{|n\lambda\rangle\}$ and $\{|n\lambda\rangle'\}$ constitute a possible set of instantaneous solutions of (1) and the freedom in choice of either one or another manifests the *gauge freedom*. The corresponding group which is used to formulate the condition of the gauge freedom is called the *gauge group*. In our case the gauge group is $\mathbb{U}(1)$.
The time evolution which we want to consider is realized via a certain time-dependence of λ, which goes along a given curve $\mathcal{C}$ in M: $t \in [0,T] \to \lambda(t) \in \mathcal{C}$. We will assume in the following that T is the period of $H(t) = H(\lambda(t))$, i.e. $\lambda(0) = \lambda(T)$, $H(0) = H(T)$, $\varepsilon_n(0) = \varepsilon_n(T)$. We will also assume for simplicity that closed path $\mathcal{C}$ completely lies in a single patch, i.e. we can choose $|n\lambda\rangle$ smoothly and uniquely on $\mathcal{C}$, $|n\lambda(0)\rangle = |n\lambda(T)\rangle$. We seek for solutions of the Schrödinger equation:

$$i\frac{\partial\psi(t)}{\partial t} = H(\lambda(t))\psi(t), \quad \psi(t) \in \mathcal{H}. \tag{3}$$

Let us also assume that at time $t = 0$, $\psi(0) = |n\lambda(0)\rangle$ for certain n. We are also particularly interested in solutions of (3) which are periodic in time in the sense that for a certain τ

$$|\psi(0)\rangle\,\langle\psi(0)| = |\psi(\tau)\rangle\,\langle\psi(\tau)|\,. \tag{4}$$

Generally speaking, depending on the speed with which λ is changed in time, τ can be arbitrary. Since the problem of finding solutions of Schrödinger equation for arbitrary τ is very broad, we restrict ourselves to the case of $\tau = T$.

Consider first the case when $[H(t), H(t')] = 0$ everywhere on path $\mathcal{C}$. Constant in time Hamiltonian obviously belongs to this class. Using the spectral resolution of the time-evolution operator in this case [2], we can show that

$$\psi(t) = e^{-i\int_0^t \varepsilon_n(\tau)\,d\tau}\psi(0) = e^{-i\alpha_{\rm dyn}(t)}\psi(0) = e^{-i\alpha_{\rm dyn}(t)}\,|n\lambda(0)\rangle\,, \tag{5}$$

where with $\alpha_{\rm dyn}(t)$ we denoted the *dynamical phase*: $\alpha_{\rm dyn}(t) = \int_0^t \varepsilon_n(\tau)\,d\tau$. The solution $\psi(t)$ from above is obviously *stationary*:

$$W(t) = |\psi(t)\rangle\,\langle\psi(t)| = |\psi(0)\rangle\,\langle\psi(0)| = W(0) = \Lambda_n(0), \tag{6}$$

where $\Lambda_n(\lambda) = |n\lambda\rangle\,\langle n\lambda|$.
Next we introduce an important less stringent assumption which is called *adiabatic approximation*. Within this approximation we assume that:

$$W(t) = \Lambda_n(t) \quad \Longleftrightarrow \quad |\psi(t)\rangle = e^{-i\alpha_\psi(t)}\,|n\lambda(t)\rangle\,, \quad t \in [0, T]. \tag{7}$$

Let us see how reasonable this assumption for *adiabatic time-evolution* is. For this we rewrite the general solution of the Schrödinger equation in the following form:

$$|\psi(t)\rangle = \sum_n c_n(t)\,|n\lambda(t)\rangle\,, \tag{8}$$

and substitute it into (3). This gives the following equations for the coefficients:

$$\partial_t c_n(t) = -i\varepsilon_n(t)c_n(t) - \sum_m c_m(t)\,\langle n\lambda(t)|\partial_t|m\lambda(t)\rangle\,. \tag{9}$$

Thus, if we start with a certain $c_n(0) = 1, c_m(0) = 0, m \neq n$, the adiabatic assumption is approximately fulfilled when $\langle n\lambda(t)\,|\,\partial_t\,|\,m\lambda(t)\rangle$ are small for $m \neq n$. One can use Eq. (1) to show that the latter matrix element can be expressed like this:

$$\langle n\lambda(t)|\partial_t|m\lambda(t)\rangle = \frac{\langle n\lambda(t)|\,\partial_t H(t)\,|m\lambda(t)\rangle}{\varepsilon_m(t) - \varepsilon_n(t)} = T^{-1}_{mn}, \tag{10}$$

which corresponds to the frequency of transition between the states $|n\lambda(t)\rangle$ and $|m\lambda(t)\rangle$. It is therefore for the states which are well-separated from each other in energy and for slowly varying (slower than the intrinsic time-scale of quantum transitions between states) in time Hamiltonians that the adiabatic approximation is more valid. It is important to remember, however, that the adiabatic wavefunction given by Eq. (7) can never be a solution of the Eq. (3) and can serve only as an approximation to it.
Under adiabatic assumption we can directly solve Eq. (9) for $c_n(t)$ by integration:

$$c_n(t) = e^{-i\int_0^t \varepsilon_n(\tau)\,d\tau}\,e^{i\int_0^t i\langle n\lambda(\tau)|\partial_\tau|n\lambda(\tau)\rangle\,d\tau} = e^{-i\alpha_{\rm dyn}}\,e^{i\gamma_n(t)}, \tag{11}$$

$$\psi(t) = e^{-i\alpha_{\rm dyn}(t)}\,e^{i\gamma_n(t)}\,|n\lambda(t)\rangle\,, \tag{12}$$

where in addition to the dynamical phase the wavefunction acquires the so-called *geometric*, or, *Berry* phase factor with the geometric phase γ_n:

$$\gamma_n(t) = \int_0^t i\,\langle n\lambda(\tau)|\partial_\tau|n\lambda(\tau)\rangle\;d\tau \mod 2\pi. \tag{13}$$

2.2 Connection and curvature

The expression for the geometric phase (13) can be easily recast into a time-independent purely geometrical representation:

$$\gamma_n(t) = \int_0^t i \langle n\lambda(\tau)|\partial_\tau|n\lambda(\tau)\rangle \, d\tau = \int_{\lambda(0)}^{\lambda(t)} i \langle n\lambda(\tau)|\partial_{\lambda_i}|n\lambda(\tau)\rangle \, d\lambda_i = \int_{\lambda(0)}^{\lambda(t)} \mathcal{A}^n, \qquad (14)$$

where $\mathcal{A}^n$ is the so-called *Berry connection* = *connection* = *connection form*:

$$\mathcal{A}^n = i \langle n\lambda|\partial_{\lambda_i}|n\lambda\rangle \, d\lambda_i = \boldsymbol{\mathcal{A}}^n \, d\boldsymbol{\lambda}, \quad \boldsymbol{\mathcal{A}}^n_i = i \langle n\lambda|\partial_{\lambda_i}|n\lambda\rangle \, . \qquad (15)$$

It can be easily shown that $\mathcal{A}^n$, and so is the geometric phase, are purely real quantities. Since function $|n\lambda\rangle$ is smooth and unique on a given patch, $\mathcal{A}^n$ is also a smooth and unique function on the corresponding patch of M. At the boundary between two patches where $|n\lambda\rangle$ and $|n\lambda\rangle'$ are related by a gauge transformation (2), the corresponding connections are related via the gauge transformation:

$$\mathcal{A}'^n = \mathcal{A}^n - d\zeta_n, \quad \boldsymbol{\mathcal{A}}'^n = \boldsymbol{\mathcal{A}}^n - \nabla\zeta_n. \qquad (16)$$

Also upon a change in the gauge of the $|n\lambda\rangle$ on the patch itself the connection transforms according to (16). Mathematically [3], the family of patches $\{O\}$ from M, together with the corresponding family of connections $\{\mathcal{A}^n\}$ and $\mathbb{U}_1$-gauge transformations between the patches $\{\zeta_n\}$ endows M with the geometric structure suitable for studies of many problems in quantum physics. The mathematical name of this structure is a $\mathbb{U}_1$ *principle fibre bundle*. The quantum theory formulated in terms of a principle fibre bundle theory with a certain group $\mathbb{G}$ of gauge transformations is often called a $\mathbb{G}$ *gauge theory*. For example, theory of electromagnetism can be elegantly recast in terms of the fibre bundle theory, with the role of connection played by the electro-magnetic vector potential $\mathbf{A}$ [3]. As we shall see, the analogy between the Berry phase theory and electromagnetism goes much deeper than sharing the name for corresponding geometrical structure.

We now consider a closed path $\mathcal{C}$ with $\lambda(0) = \lambda(T)$ and $\gamma_n(\mathcal{C}) := \gamma_n(T)$. Let us trace how $\gamma_n(\mathcal{C})$ changes upon an arbitrary smooth and unique change of gauge in the patch in which $\mathcal{C}$ lies:

$$\gamma'_n(\mathcal{C}) = i \oint_{\mathcal{C}} \mathcal{A}'^n = i \oint_{\mathcal{C}} \mathcal{A}^n - i \oint_{\mathcal{C}} d\zeta_n = i \oint_{\mathcal{C}} \mathcal{A}^n = \gamma_n(\mathcal{C}), \qquad (17)$$

since $\zeta_n(0) = \zeta_n(T)$. This means that the Berry phase of closed path $\mathcal{C}$ is an intrinsically *gauge-invariant* property, which is another manifestation of its purely geometrical meaning. One can show that $\gamma_n(\mathcal{C})$ remains constant even under action of a more general class of gauge transformation, and cannot be "gauged away". While it cannot be done for the whole path, on a part of $\mathcal{C}$ the Berry phase can be indeed gauged away by choosing the so-called *parallel transport gauge* realized by smooth and unique functions ζ_n^{PT}, defined only on the part of $\mathcal{C}$ which we denote as $\tilde{\mathcal{C}}$. Given the initial choice of $\{|n\lambda\rangle\}$, ζ_n^{PT} functions satisfy the following condition:

$$\boldsymbol{\mathcal{A}}^{\mathrm{PT,n}} = \boldsymbol{\mathcal{A}}^n - \nabla\zeta_n^{\mathrm{PT}} = 0, \qquad (18)$$

equivalent to the condition that

$$\langle n\lambda^{\mathrm{PT}}|\nabla_\lambda|n\lambda^{\mathrm{PT}}\rangle = 0, \quad \gamma_n^{\mathrm{PT}}(\tilde{\mathcal{C}}) := i \int_{\tilde{\mathcal{C}}} \mathcal{A}^{\mathrm{PT},n} = 0. \qquad (19)$$

Besides being very convenient, the parallel transport gauge is also easy to constuct. Namely, given a certain starting point λ and a set of starting $\{|n\lambda\rangle^{\rm PT}\}$ at this point, a set of "parallel transported" instantaneous states at an infinitesimally to λ close point $\lambda+\epsilon$ can be constucted using perturbation theory:

$$|n,\lambda+\epsilon\rangle^{\rm PT} = |n\lambda\rangle^{\rm PT} + \sum_{m\neq n} \frac{\langle m\lambda^{\rm PT}|H(\lambda+\epsilon)-H(\lambda)|n\lambda^{\rm PT}\rangle}{\varepsilon_n(\lambda)-\varepsilon_m(\lambda)} \left|m\lambda^{\rm PT}\right\rangle. \tag{20}$$

It can be easily checked that such constructed instantaneous solutions are indeed "parallel transported". It is common to use the parallel transport gauge for evaluation of the quantities which are gauge-invariant *locally* for each point λ on M. A most important example of such quantity is the (Berry) curvature.

(Berry) curvature is defined as a rank-2 antisymmetric tensor with the components given by:

$$\Omega^n_{ij} := \partial_{\lambda_i}\mathcal{A}^n_j - \partial_{\lambda_j}\mathcal{A}^n_i = -2{\rm Im}\left\langle \frac{\partial}{\partial_{\lambda_i}} n\lambda \middle| \frac{\partial}{\partial_{\lambda_j}} n\lambda \right\rangle. \tag{21}$$

It is easy to see that the components of Ω^n are *locally* gauge-invariant with respect to the gauge transformations (16). In the language of differential forms, Ω^n is a 2-form given by $\Omega^n = d\mathcal{A}^n$. Using the Stokes' theorem the expression for the Berry phase (14) can be rewritten as:

$$\gamma_n(\mathcal{C}) = i\oint_{\mathcal{C}} \mathcal{A}^n = \int_S \Omega^n, \tag{22}$$

where S is a "surface" in M which $\mathcal{C}$ encompasses: $\partial S = \mathcal{C}$. The Berry phase thus equals to the *flux* of the Berry curvature through S. A very useful expression for the curvature can be obtained within the parallel transport gauge (20), in which the perturbation theory expression for the derivative of the $|n\lambda\rangle$ entering (21) can be written down (we omit the "PT" label for simplicity):

$$\partial_{\lambda_i}|n\lambda\rangle = \sum_{m\neq n} \frac{\langle m\lambda|\partial_{\lambda_i}H(\lambda)|n\lambda\rangle}{\varepsilon_n(\lambda)-\varepsilon_m(\lambda)}|m\lambda\rangle, \tag{23}$$

which gives the following gauge-invariant expression for the curvature:

$$\Omega^n_{ij}(\lambda) = -2{\rm Im}\sum_{m\neq n} \frac{\langle m\lambda|\partial_{\lambda_i}H(\lambda)|n\lambda\rangle\langle n\lambda|\partial_{\lambda_j}H(\lambda)|m\lambda\rangle}{[\varepsilon_n(\lambda)-\varepsilon_m(\lambda)]^2}. \tag{24}$$

The latter expression is not only useful practically, but it is also rather instructive since it underlines the role of the band degeneracies as the sources of the Berry curvature around them. We will see several examples of connection between the band degeneracies and the curvature in the course of this manuscript.

In case of M being a part of $\mathbb{R}^3$ the curvature tensor can be seen as a vector $\mathbf{\Omega}^n$ in $\mathbb{R}^3$, with components given by:

$$\mathbf{\Omega}^n_i := (1/2)\epsilon_{ijk}\Omega^n_{jk}, \tag{25}$$

and the relation (21) between the connection and curvature becomes:

$$\mathbf{\Omega}^n = {\rm curl}\boldsymbol{\mathcal{A}}^n. \tag{26}$$

There can be cases when the connection is the so-called "pure gauge" connection, i.e., there exists a smooth and unique vector field $\mathbf{f}$ on M, so that $\boldsymbol{\mathcal{A}}^n = \nabla\mathbf{f}$. Clearly, in this case the

Berry curvature equals zero identically on M and the Berry phase is vanishing for any given closed path in M. In the rest of the manuscript we are particularly interested in situations for which this does not happen.

Finally, we would like to remark on the occurrence of the so-called *non-Abelian* Berry phase. It differs from the *Abelian* case considered above in that one has to deal with an N-fold degeneracy of eigenvalue ε_n at any point λ from M. Given a certain set of instantaneous solutions $\{|n\alpha\lambda\rangle\}_{\alpha=1}^{N}$ which correspond to ε_n for each λ we can construct corresponding eigenprojectors $\Lambda_n(\lambda) = \sum_{\alpha=1}^{N} |n\alpha\lambda\rangle \langle n\alpha\lambda|$. The eigenprojectors are constant with respect to the gauge transformations realized by the unitary transformations $\mathcal{U}^N$ from $\mathbb{U}(N)$:

$$|n\vec{\alpha}\lambda\rangle' = \mathcal{U}^N |n\vec{\alpha}\lambda\rangle, \tag{27}$$

where $|n\vec{\alpha}\lambda\rangle$ is a vector consisting of ordered instantaneous solutions. The adiabatic assumption for the non-Abelian case can be formulated in analogy to the Abelian case: the solution of the Schrödinger equation has to reside in n'th eigenspace, $|\psi(t)\rangle \langle\psi(t)| = \Lambda_n(\lambda(t))$. The most general shape of the wavefunction which solves (3) under the adiabatic assumption reads:

$$|\psi(t)\rangle = \sum_{\alpha} c_\alpha^n(t) \, |n\alpha\lambda(t)\rangle \, . \tag{28}$$

Substituting the latter expression into (3) gives us a system of equations for the c-coefficients, which can be solved in analogy to the Abelian case yielding:

$$\vec{c^n}(t) = \mathcal{T} \exp\left(\int_0^t d\tau \left\{ -i\varepsilon_n \mathbb{I} + i\mathcal{A}_N^n(\tau) \right\} \right) \vec{c^n}(0), \tag{29}$$

where $\mathcal{T}$ is the time-ordering operator, $\mathbb{I}$ is the $N \times N$ identity matrix, and $\mathcal{A}_N^n$ is a Hermitian-matrix valued 1-form, given by components:

$$[\mathcal{A}_N^n(\lambda(t))]^{\alpha\beta} := i \, \langle n\alpha\lambda(t)|\partial_t|n\beta\lambda(t)\rangle \, dt. \tag{30}$$

The form $\mathcal{A}_N^n$ is called the *non-Abelian (Berry) connection*, in analogy to the Abelian connection. The gauge transformation of the non-Abelian connection can be deduced from (27):

$$\mathcal{A}_N'^n = (\mathcal{U}^N)^{-1} \mathcal{A}_N^n \mathcal{U}^N + i(\mathcal{U}^N)^{-1} d\mathcal{U}^N. \tag{31}$$

Obviously this relation reduces to (16) if $N = 1$. Equation (29) allows us to calculate the total "phase" a wavefunction acquires during adiabatic evolution with the period T along a closed path $\mathcal{C}$, i.e., $\psi(T) = \mathcal{U}_\psi \psi(0)$, with the "phase" $\mathcal{U}_\psi$ given by:

$$\mathcal{U}_\psi = \exp\left(-i \int_0^T \varepsilon_n(\lambda(t)) dt \right) \mathcal{P} \exp\left(i \oint_{\mathcal{C}} \mathcal{A}_N^n \right), \tag{32}$$

where the first factor is the analogon of the Abelian dynamical phase, while the second matrix is the *non-Abelian Berry phase*. While having most of the geometric "niceties" of the Abelian Berry phase, the non-Abelian phase is a matrix with elements which are not separately gauge-invariant. Most relevant are thus the trace (also known as the Wilson loop) and the eigenvalues of the non-Abelian Berry phase, which are indeed gauge-invariant quantities. The expression for the matrix valued non-Abelian Berry curvature reads:

$$\Omega_{N,ij}^n := \partial_{\lambda_i} \mathcal{A}_{N,j}^n - \partial_{\lambda_j} \mathcal{A}_{N,i}^n + [\mathcal{A}_{N,i}^n, \mathcal{A}_{N,j}^n], \tag{33}$$

which reduces to (21) for $N = 1$.

2.3 Topological phase: Aharonov-Bohm effect

The Berry phase which we introduced is *geometric* in nature, i.e., it depends on the local geometry of the path $\mathcal{C}$. This also means that when the path $\mathcal{C}$ is changed, the geometric phase will change as well. There can be situations, however, when the Berry phase is not modified when the path $\mathcal{C}$ is subject to smooth transformations, in which case the Berry phase is *topological* in nature. A very prominent example of a topological phase arises within the *Aharonov-Bohm effect* (AB-effect).
Imagine an infinitesimally small cylinder D with the axis along z and cutting through the center of coordinates in $\mathbb{R}^3$. Inside the cylinder only we generate a constant magnetic field $\mathbf{B}$ along the cylinder axis and a magnetic flux through the cylinder's cross section of Φ. Consider an electron confined to a box which is positioned at point $\mathbf{R}$ which lies very far away from the cylinder. The Hamiltonian of our electron with respect to the center of the box reads:

$$H = H(\mathbf{p} - \mathbf{A}(\mathbf{x}), \mathbf{x} - \mathbf{R}), \tag{34}$$

where $\mathbf{A}$ is the electromagnetic vector potential with $\mathrm{curl}\mathbf{A} = \mathbf{B}$. We can thus look at the dependence of the Hamitonian on $\mathbf{R}$ as a parametric dependence considered previously with $\lambda = \mathbf{R}$. According to the theory above, we have to find the eigenvalues and eigenvectors of H:

$$H(\mathbf{p} - \mathbf{A}(\mathbf{x}), \mathbf{x} - \mathbf{R}) \, |n\mathbf{R}\rangle = \varepsilon_n(\mathbf{R}) \, |n\mathbf{R}\rangle \, . \tag{35}$$

It is easy to see that to find $\{|n\mathbf{R}\rangle\}$ we can use the eigenvalues $\varepsilon_n(\mathbf{R})$ and eigenfunctions ψ_n^0 of the free Hamiltonian $H_0 = H(\mathbf{p}, \mathbf{x} - \mathbf{R})$:

$$\langle \mathbf{x} \, | \, n\mathbf{R} \rangle = e^{i \int_{\mathbf{R}}^{\mathbf{x}} \mathbf{A} \, d\mathbf{l}} \, \psi_n^0(\mathbf{x} - \mathbf{R}), \tag{36}$$

where the path between $\mathbf{x}$ and $\mathbf{R}$ cannot pass through D. We can thus readily show that the Berry connection $\boldsymbol{\mathcal{A}}_n$ is given by:

$$\begin{aligned} \boldsymbol{\mathcal{A}}^n = i \, \langle n\mathbf{R} \, | \, \nabla_{\mathbf{R}} \, | \, n\mathbf{R} \rangle &= i \int_{\mathbb{R}^3} d\mathbf{x} \, \psi_n^{0*}(\mathbf{x} - \mathbf{R}) \left\{ -i\mathbf{A}(\mathbf{R}) \psi_n^0(\mathbf{x} - \mathbf{R}) + \nabla_{\mathbf{R}} \psi_n^0(\mathbf{x} - \mathbf{R}) \right\} \\ &= \mathbf{A}. \end{aligned}$$

In turn the curvature is given by:

$$\boldsymbol{\Omega}^n = \mathrm{curl}\boldsymbol{\mathcal{A}} = \mathrm{curl}\mathbf{A} = \mathbf{B}. \tag{37}$$

This means that for our problem the Berry phase $\gamma_n(\mathcal{C})$ acquired during adiabatic motion of our electron box around the cylinder with the magnetic field along a path $\mathcal{C}$ is given by:

$$\gamma_n(\mathcal{C}) = \oint_{\mathcal{C}} \mathbf{A} \, d\mathbf{l} = -\Phi. \tag{38}$$

Note that γ_n does not depend neither on the "band" index n nor on the path $\mathcal{C}$, as long as the cylinder is encircled the same number of times. This presents an elegant example of the topological phase which does not change upon smooth transformations of $\mathcal{C}$. What is very important for us is that the results of Eq. (37) can be generalized to the case of electrons not necessarily confined to a box, with the magnetic field $\mathbf{B}$ present everywhere in $\mathbb{R}^3$ and possibly having non-trivial spatial distribution. This will result in the Berry phase restoring back its geometrical rather than topological meaning, since in the latter case:

$$\gamma_n(\mathcal{C}) = \oint_{\mathcal{C}} \mathbf{A} \, d\mathbf{l} = -\Phi(\mathcal{C}), \tag{39}$$

with $\Phi(\mathcal{C})$ as the $\mathcal{C}$-dependent flux of the magnetic field through the area enclosed by $\mathcal{C}$.

2.4 Dirac monopole and spin-$\frac{1}{2}$ problem

As an example of the adiabatic dynamics we want to study the Hamiltonian of a spin-$\frac{1}{2}$ particle in magnetic field $\mathbf{B} = B\cdot\mathbf{n}$ in $\mathbb{R}^3$, with the unit vector $\mathbf{n}$ along $\mathbf{B}$, described by the Hamiltonian:

$$H(\mathbf{n}) = b_0\boldsymbol{\sigma}\cdot\mathbf{B} = b\boldsymbol{\sigma}\cdot\mathbf{n}, \tag{40}$$

where $b = b_0B$ is a constant which we would like to keep unspecified, and $\boldsymbol{\sigma}$ is the vector of Pauli matrices. The parameter space of our problem is the 2-sphere S^2 embedded in $\mathbb{R}^3$, with the parameter $\mathbf{n}$ playing the role of λ. We parametrize $\mathbf{n}$ in spherical coordinates: $\mathbf{n} = (\sin\theta\cos\varphi, \sin\theta\cos\varphi, \cos\theta)$, where θ and φ run between 0 and π, and 0 and 2π, respectively, see Fig. 1. For each of the values of $\mathbf{n}$ we have to find the instantaneous eigenstates and eigenvalues of $b\boldsymbol{\sigma}\cdot\mathbf{n}$. There are two eigenvalues for each $\mathbf{n}$: $\varepsilon_+ = +b$ and $\varepsilon_- = -b$, and two corresponding eigenstates of the projection of $\boldsymbol{\sigma}$ onto $\mathbf{n}$: $|+\mathbf{n}\rangle$ and $|-\mathbf{n}\rangle$, or, alternatively:

$$H(\mathbf{n})\,|\sigma\mathbf{n}\rangle = b\sigma\,|\sigma\mathbf{n}\rangle\,, \tag{41}$$

where σ takes values of $+/-$ or $+1/-1$ depending on the situation. Note that the eigenvalues do not depend on $\mathbf{n}$. More generally, the operator $\boldsymbol{\sigma}$ in (40) can be replaced with $2\mathbf{J}$, where $\mathbf{J}$ is the angular momentum operator of our quantum system. In this case σ is not anymore restricted to take only two values, and the problem can be reduced to more than two copies of the problem for spin-up and spin-down states only. Nevertheless the conclusions and derived expressions remain basically the same as those for the spin-$\frac{1}{2}$ problem, which we are to consider below.
Our task now is to choose as smooth a gauge of $|+\mathbf{n}\rangle$ and $|-\mathbf{n}\rangle$ as we can. We note that the vector $\mathbf{n}$ can be obtained from $\mathbf{e}_3$ via the following $\mathrm{SO}(3)$ rotation: $\mathcal{R}(\theta,\phi) = \mathcal{R}_3(\varphi)\mathcal{R}_2(\theta)\mathcal{R}_3(-\varphi)$, where $\mathcal{R}_3(\varphi)$ performs rotation around $\mathbf{e}_3$ by angle φ, and $\mathcal{R}_2(\theta)$ is a rotation around $\mathbf{e}_2$ by θ. The corresponding rotation of the spinor wavefunction is an element of $\mathrm{SU}(2)$ and is given by:

$$U(\theta,\varphi) = U_3(\varphi)U_2(\theta)U_3(-\varphi) = e^{-i\varphi\sigma_z/2}\,e^{-i\theta\sigma_y/2}\,e^{i\varphi\sigma_z/2}. \tag{42}$$

We thus define:

$$|+\mathbf{n}\rangle = U(\theta,\varphi)\,|+\mathbf{e}_3\rangle\,,\quad |-\mathbf{n}\rangle = U(\theta,\varphi)\,|-\mathbf{e}_3\rangle \quad\Leftrightarrow\quad |\sigma\mathbf{n}\rangle = U(\theta,\varphi)\,|\sigma\mathbf{e}_3\rangle\,. \tag{43}$$

It can be easily shown that such choice for $|+\mathbf{n}\rangle$ and $|-\mathbf{n}\rangle$ is smooth everywhere on S^2 except for the south pole S corresponding to vector $-\mathbf{e}_3$. To achieve smoothness at S, we need to perform a gauge transformation into a new basis:

$$|\sigma\mathbf{n}\rangle' = |\sigma\mathbf{n}'\rangle = e^{-i\sigma\varphi}\,|\sigma\mathbf{n}\rangle\,,\quad \zeta_\sigma(\mathbf{n}) = -\sigma\varphi \tag{44}$$

and it can be shown that the primed instantaneous eigenstates are smooth everywhere on S^2 except for the north pole N which corresponds to $\mathbf{e}_3$. It is a property of the spin-$\frac{1}{2}$ system in magnetic field that enforces us to introduce at least two patches on S^2 where the instantaneous eigenstates are smooth, and a phase "twist" = gauge transformation between two smooth families at the overlap between the two patches. We denote these patches by $O_1 = S^2 - \mathrm{S}$ and $O_2 = S^2 - \mathrm{N}$. Note that we could also freely choose for O_1 and O_2 the northern and southern hemispheres including the equator. In this case the gauge transformation would act only at the equator (to be rigorous, since the patches have to be open sets, by equator here we mean an infinitesimally small open "belt" which includes the equator).

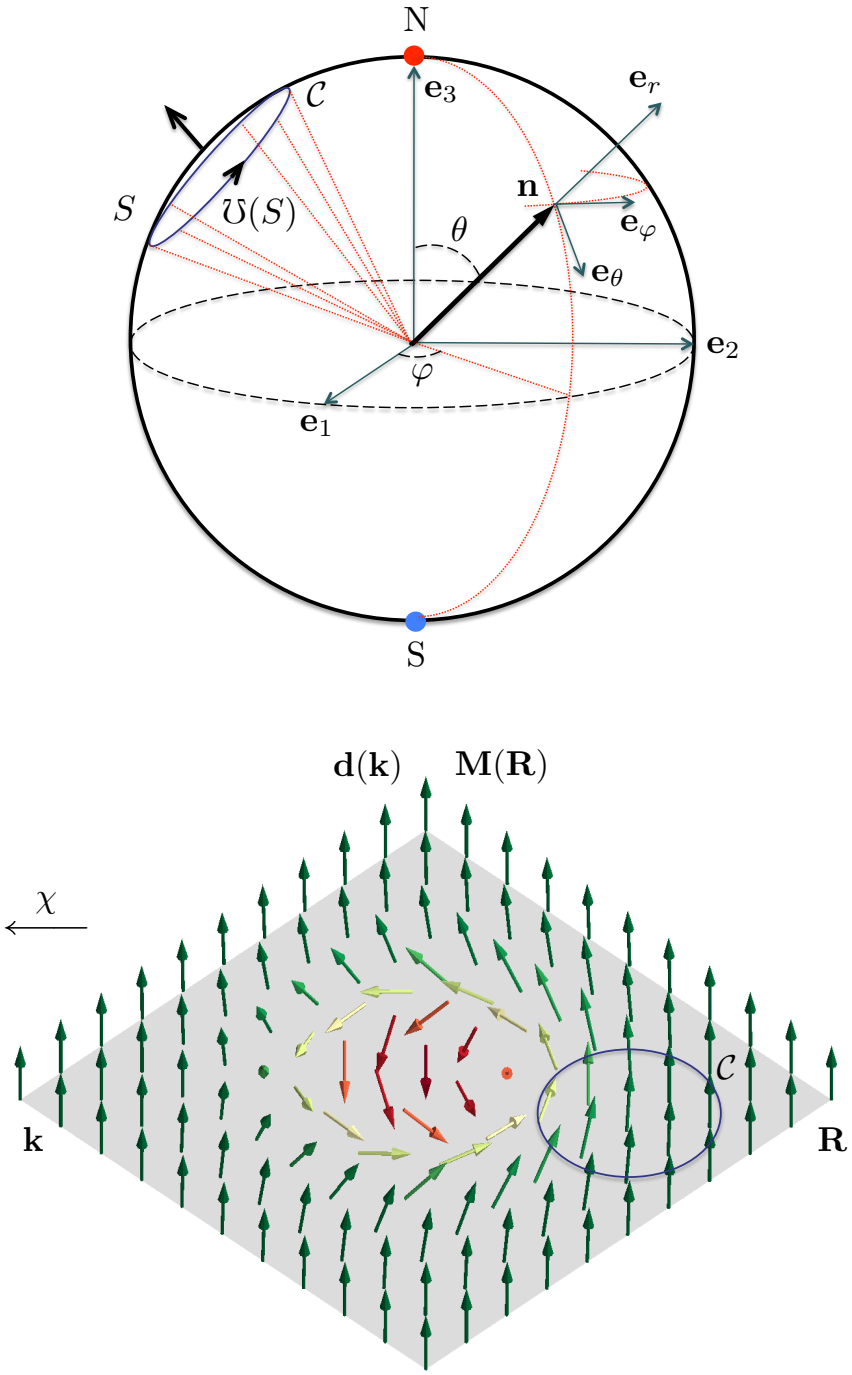


Fig. 1: *Top: geometry of spin-$\frac{1}{2}$ in magnetic field. Bottom: skyrmion distribution of vector field* $\mathbf{d}(\mathbf{k})$ *as a function of the Bloch vector* $\mathbf{k}$ *in Eq. (112); or of the magnetization field* $\mathbf{M}(\mathbf{R})$ *as a function of real-space coordinate* $\mathbf{R}$ *for magnetic texture as given by Eq. (154). The transformation from the* $\mathbf{k}(\mathbf{R})$*-space into the space of the magnetic field from spin-$\frac{1}{2}$ problem is realized by a map* χ*, which allows to relate the Berry phase of path* $\mathcal{C}$ *in the lower figure to the Berry phase of the corresponding path* $\mathcal{C}$ *in the upper figure. The direction of* $\mathbf{d}(\mathbf{R})$ *in the lower plot spans the whole sphere in the upper plot once. The artist's view of the skyrmion is by Jürgen Weischenberg.*

We are now ready to define the connections on O_1 and O_2. They are given by corresponding components:

$$\mathcal{A}^{\sigma}_{\theta}(\theta,\varphi) = \langle \sigma\mathbf{n} \,|\, \partial_\theta \,|\, \sigma\mathbf{n}\rangle = 0, \quad \mathcal{A}^{\sigma}_{\varphi}(\theta,\varphi) = \langle \sigma\mathbf{n} \,|\, \partial_\varphi \,|\, \sigma\mathbf{n}\rangle = -(1/2)\sigma(1-\cos\theta), \tag{45}$$

$$\mathcal{A}'^{\sigma}_{\theta}(\theta,\varphi) = \langle \sigma\mathbf{n}' \,|\, \partial_\theta \,|\, \sigma\mathbf{n}'\rangle = 0, \quad \mathcal{A}'^{\sigma}_{\varphi}(\theta,\varphi) = \langle \sigma\mathbf{n}' \,|\, \partial_\varphi \,|\, \sigma\mathbf{n}'\rangle = (1/2)\sigma(1+\cos\theta). \tag{46}$$

Indeed, according to the gauge transformation rule, $\mathcal{A}^{\sigma}_{\varphi} - \mathcal{A}'^{\sigma}_{\varphi} = -\sigma = \partial_\varphi \zeta_\sigma$. The only component of the curvature can be also readily evaluated:

$$\Omega^{\sigma}_{\theta\varphi}(\theta,\varphi) = -\sigma\sin\theta/2. \tag{47}$$

Since our sphere S^2 is naturally embedded into $\mathbb{R}^3$, the connection and curvature can be recast as vectors in $\mathbb{R}^3$:

$$\boldsymbol{\mathcal{A}}^{\sigma} = \hat{\mathcal{A}}^{\sigma}_{r}\mathbf{e}_r + \hat{\mathcal{A}}^{\sigma}_{\theta}\mathbf{e}_\theta + \hat{\mathcal{A}}^{\sigma}_{\varphi}\mathbf{e}_\varphi = \hat{\mathcal{A}}^{\sigma}_{\varphi}\mathbf{e}_\varphi, \quad \boldsymbol{\mathcal{A}}'^{\sigma} = \hat{\mathcal{A}}'^{\sigma}_{\varphi}\mathbf{e}_\varphi, \quad \boldsymbol{\Omega}^{\sigma} = \Omega^{\sigma}_{r}\,\mathbf{e}_r, \tag{48}$$

where the components read:

$$\hat{\mathcal{A}}^{\sigma}_{\varphi} = \frac{\sigma(\cos\theta-1)}{2r\sin\theta}, \quad \hat{\mathcal{A}}'^{\sigma}_{\varphi} = \frac{\sigma(\cos\theta+1)}{2r\sin\theta}, \quad \Omega^{\sigma}_{r} = -\frac{\sigma}{2r^2}. \tag{49}$$

The latter expressions are remarkable in that they present a realization of so-called *Dirac monopole*. Dirac monopole corresponds to a problem of a single *monopole* magnetic charge of magnitude g at the origin which provides the source of the magnetic field $\mathbf{B}_m$:

$$\mathrm{div}\mathbf{B}_m = 4\pi g\,\delta^3(\mathbf{r}), \tag{50}$$

with the solution of this equation being

$$\mathbf{B}_m = \frac{g}{r^2}\,\mathbf{e}_r, \tag{51}$$

which is exactly the expression for the curvature with $g = -\sigma/2$. The flux of this magnetic field through S^2 is $4\pi g$. Obviously, the magnetic field $\mathbf{B}_m$ and $\boldsymbol{\Omega}^{\sigma}$ have a singularity at $r = 0$. As discussed previously, this singularity in the Berry curvature is due to a degeneracy between ε_+ and ε_- at the origin. In this sense, we can say that such degeneracy points in the spectrum serve as "sources" of the Berry curvature everywhere around them, in the same way that the magnetic monopole g serves as a source of the magnetic field $\mathbf{B}_m$. It is important to remember that it is the Hamiltonian which we started with which serves as a source of non-trivial geometry of our problem, i.e., the fact that $\sigma \neq 0$ gives rise to the non-trivial gauge transformation which in turn leads to the non-zero monopole charge at the origin. However, as we shall see, the geometry of the problem can also have important consequences for the Hamiltonian and we will show that σ *has to* be quantized, even though such assumption was not necessary to make from the beginning.

To show this, lets evaluate the Berry phase of a certain path $\mathcal{C}$ which our spin "draws" on S^2 as it follows the adiabatically slow magnetic field. If we suppose that $\mathcal{C}$ does not include the south pole, then we can easily write the Berry phase as:

$$\gamma_\sigma(\mathcal{C}) = \oint_{\mathcal{C}} \boldsymbol{\mathcal{A}}^{\sigma}\, d\mathbf{l} = \int_S \boldsymbol{\Omega}^{\sigma}\, d\mathbf{S} = -\sigma\mho(S)/2 \mod 2\pi, \tag{52}$$

where $\mathcal{C} = \partial S$, $\mho(S)$ is the solid angle of surface S, and the orientations of $d\mathbf{l}$ and $d\mathbf{S}$ are shown in Fig. 1. On the other hand, we can use surface $S' = S^2 - S$ to perform the integral of the curvature:

$$\gamma_\sigma(\mathcal{C}) = -\sigma\mho(S)/2 = -\int_{S'} \mathbf{\Omega}^\sigma \, d\mathbf{S} = \sigma\mho(S')/2 = \sigma(4\pi - \mho(S))/2 \mod 2\pi, \tag{53}$$

and thus we independently arrive at the *quantization* of σ:

$$\sigma \in \mathbb{Z}, \tag{54}$$

which is the sole property of the geometry of our problem. This quantization condition leads us to the definition of the *(first) Chern number*:

$$C_\sigma = \frac{1}{2\pi}\int_{S^2} \mathbf{\Omega}^\sigma \, d\mathbf{S} = -\sigma, \tag{55}$$

which stands for the flux of the Berry curvature through the whole sphere. It can be shown that for "well-behaved" two-dimensional compact manifolds for which we will also assume that they have no boundary (e.g. sphere, torus) the Chern number is a topological invariant of the manifold with the geometry on it specified by a family of connections $\{\mathcal{A}\}$ and gauge transformations $\{\zeta\}$ belonging to the gauge group (i.e. of the so-called mathematically *fibre bundle*), and that $C \in \mathbb{Z}$. Intuitively, it is clear why the manifold has to have no boundary, i.e. to be closed. Imagine that instead of S^2 we have only a part of it. Then we can easily imagine how by a smooth deformation we could "pull out" the Dirac monopole out of the sphere, thus continuously changing the flux of the curvature through the part of S^2. On the other hand, smooth transformations can only change the position of the monopole inside the whole sphere while keeping the value of the flux quantized. Our example of spin-$\frac{1}{2}$ particle in magnetic field and proof of quantization of C_σ is a particular example of this general mathematical fact. As we have also seen, the non-zeroness of the Chern number is intrinsically related to our inability of choosing a smooth gauge over the *entire* manifold. We refer to [3] for further discussions on higher-dimensional manifolds, theory of Chern numbers and characteristic classes.

3 Selected Applications of Geometric Phase in Condensed Matter

3.1 Berry curvature for Bloch electrons

For electrons in a periodic solid the eigenstates $\psi_{n\mathbf{k}}$ of the Hamiltonian can be classified by quantum numbers $(\mathbf{k}, n)$, where $\mathbf{k}$ lies in the so-called Brillouin zone (BZ), and n is a discrete index numbering the bands. The eigenfunctions can be written in the following form:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \tag{56}$$

where $u_{n\mathbf{k}}$ has the periodicity of the lattice. It then follows that instead of writing

$$H\psi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r}) \tag{57}$$

we can write

$$H(\mathbf{k})u_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}}u_{n\mathbf{k}}(\mathbf{r}), \tag{58}$$

where

$$H(\mathbf{k}) \equiv H_{\mathbf{k}} \equiv e^{-i\mathbf{k}\cdot\mathbf{r}} \cdot H \cdot e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{59}$$

We were thus able to rewrite the problem (57) in terms of an eigenvalue problem of a $\mathbf{k}$-dependent Hamiltonian, acting on the same Hilbert space of periodic functions for every $\mathbf{k}$. This is exactly the setup suitable for studies of the Berry phase effects, if we identify the parameter λ from general mathematical theory with the Bloch vector $\mathbf{k}$. The corresponding (Berry) connection of non-degenerate band n according to (15) then reads:

$$\mathcal{A}^n(\mathbf{k}) = i \left\langle u_{n\mathbf{k}} \,|\, \partial_{\mathbf{k}} u_{n\mathbf{k}} \right\rangle, \tag{60}$$

and the components of the Berry curvature tensor of band n are given by:

$$\Omega^n_{ij}(\mathbf{k}) = -2\mathrm{Im} \left\langle \partial_{k_i} u_{n\mathbf{k}} \,\middle|\, \partial_{k_j} u_{n\mathbf{k}} \right\rangle. \tag{61}$$

The components of the Berry curvature itself are sometimes written as a vector

$$\Omega^n_i(\mathbf{k}) := \mathbf{\Omega}^n_i(\mathbf{k}) = (1/2)\epsilon_{lmi}\Omega^n_{lm}(\mathbf{k}) = -\mathrm{Im} \left\langle \partial_{\mathbf{k}} u_{n\mathbf{k}} \,|\, \times \,|\, \partial_{\mathbf{k}} u_{n\mathbf{k}} \right\rangle. \tag{62}$$

The driving force behind the dynamics of an electron residing at a certain $\mathbf{k}$-point and band n could be an external electric or magnetic field as well as dependence of the Hamiltonian on another parameter, which, in a localized picture, cause the motion of an electron along certain orbits in $\mathbf{k}$-space and $\mathbf{r}$-space, as we shall see in detail in the following. The perturbation theory expression for the $\mathbf{k}$-space Berry curvature, looking at Eq. (24), can be written as:

$$\Omega^n_{ij}(\mathbf{k}) = -2\,\mathrm{Im} \sum_{m\neq n} \frac{\left\langle u_{n\mathbf{k}} \,|\, \partial_{k_i} H_{\mathbf{k}} \,|\, u_{m\mathbf{k}} \right\rangle \left\langle u_{m\mathbf{k}} \,\middle|\, \partial_{k_j} H_{\mathbf{k}} \,\middle|\, u_{n\mathbf{k}} \right\rangle}{(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}})^2}. \tag{63}$$

We will also consider a situation, in which, besides the dependence on $\mathbf{k}$, the Hamiltonian of the system depends at the same time on another (multi-dimensional) parameter λ, that is, $H = H(\mathbf{k}, \lambda)$. Generally speaking, the Berry curvature form in this extended $(\lambda, \mathbf{k})$ space has components, which we call $\Omega^n_{\mathbf{kk}} \equiv \Omega^n_{\mathbf{k}}$ and $\Omega^n_{\lambda\lambda} \equiv \Omega^n_{\lambda}$ and which are expressed in terms of derivatives of $u^{\lambda}_{n\mathbf{k}}$ with respect to only $\mathbf{k}$ or λ, respectively, according to (21). However, there is also the component of the Berry curvature form, which involves both λ- and $\mathbf{k}$-derivatives:

$$\Omega^n_{\lambda\mathbf{k}} = -2\mathrm{Im} \left\langle \partial_{\mathbf{k}} u^{\lambda}_{n\mathbf{k}} \,\middle|\, \partial_{\lambda} u^{\lambda}_{n\mathbf{k}} \right\rangle = -2\,\mathrm{Im} \sum_{m\neq n} \frac{\left\langle u_{n\mathbf{k}} \,|\, \partial_{\mathbf{k}} H_{\mathbf{k}} \,|\, u_{m\mathbf{k}} \right\rangle \left\langle u_{m\mathbf{k}} \,|\, \partial_{\lambda} H_{\mathbf{k}} \,|\, u_{n\mathbf{k}} \right\rangle}{(\varepsilon_{n\mathbf{k}} - \varepsilon_{m\mathbf{k}})^2}. \tag{64}$$

We call this part of the Berry curvature the *mixed Berry curvature*, and we will discuss it in detail in different contexts in the following sections.

3.2 Electric polarization

The example of the electric polarization, besides being of great importance in solid state physics, will allow us to delve into the concept of extended parameter space of reciprocal $\mathbf{k}$-vectors as well as external parameter λ, with the generic $(\mathbf{k}, \lambda)$-dependence of the Hamitonian and extension of the Berry curvature to higher dimensions.

Within the Born-Oppenheimer approximation, which assumes that the motion of electrons is much faster than the slow motion of the ions, we can separate the ionic and electronic terms

in the charge density. The ionic charge density is simply given by sum of point charges at the positions of the atoms, while we concentrate further on the electronic contribution to the polarization. Within the single-particle picture the electronic part of the charge density is given by:

$$\rho^{\lambda}(\mathbf{r}) = \sum_{n \leq M} |\psi_n^{\lambda}(\mathbf{r})|^2, \tag{65}$$

where λ is an external parameter which for example specifies the atomic displacements, M is the highest occupied level and ψ_n^{λ} are the single-particle states of the *finite* crystal:

$$\left(-\frac{\nabla^2}{2} + V^{\lambda}(\mathbf{r})\right) \psi_n^{\lambda}(\mathbf{r}) = \varepsilon_n^{\lambda} \psi_n^{\lambda}(\mathbf{r}). \tag{66}$$

For a finite sample with the volume V_c the electronic part of the electric polarization is given by

$$\mathbf{P}^{\lambda} = \frac{1}{V_c} \int_{\mathbb{R}_3} \mathbf{r} \rho^{\lambda}(\mathbf{r})\, d\mathbf{r} = \frac{1}{V_c} \int_{\text{sample}} \mathbf{r} \rho^{\lambda}(\mathbf{r})\, d\mathbf{r}, \tag{67}$$

while its derivative with respect to λ expressed in terms of the matrix elements of the $\mathbf{r}$-operator is well-defined:

$$\partial_{\lambda} \mathbf{P}^{\lambda} = \frac{1}{V_c} \sum_{n \leq M} \left(\left\langle \partial_{\lambda} \psi_n^{\lambda} \,\middle|\, \mathbf{r} \,\middle|\, \psi_n^{\lambda} \right\rangle + \left\langle \psi_n^{\lambda} \,\middle|\, \mathbf{r} \,\middle|\, \partial_{\lambda} \psi_n^{\lambda} \right\rangle \right) = \frac{1}{V_c} 2 \operatorname{Re} \sum_{n \leq M} \left\langle \partial_{\lambda} \psi_n^{\lambda} \,\middle|\, \mathbf{r} \,\middle|\, \psi_n^{\lambda} \right\rangle. \tag{68}$$

Consider now such a small change in λ so that the change in the potential can be treated within the perturbation theory. In this case the derivative of the wave function $\partial_{\lambda} \psi_n^{\lambda}$ in terms of other states within the first order perturbation theory is given by Eq. (23) (we will soon see that $\partial_{\lambda} \mathbf{P}^{\lambda}$ is a gauge-invariant quantity and thus we can safely choose the parallel transport gauge), which leads to the following result:

$$\partial_{\lambda} \mathbf{P}^{\lambda} = \frac{1}{V_c} 2 \operatorname{Re} \sum_{n \leq M} \sum_{m \neq n}^{\infty} \frac{\left\langle \psi_n^{\lambda} \,\middle|\, \mathbf{r} \,\middle|\, \psi_m^{\lambda} \right\rangle \left\langle \psi_m^{\lambda} \,\middle|\, \partial_{\lambda} H^{\lambda} \,\middle|\, \psi_n^{\lambda} \right\rangle}{\varepsilon_n^{\lambda} - \varepsilon_m^{\lambda}}. \tag{69}$$

In case of a finite crystal the regions where the wavefunctions and the potential are non-zero can be considered finite, thus the matrix elements of the position operator in the expression above are well-defined. We will rewrite them now, however, using the identity valid for $m \neq n$ (again properly defined only for finite samples)

$$\left\langle \psi_n^{\lambda} \,\middle|\, \mathbf{r} \,\middle|\, \psi_m^{\lambda} \right\rangle = -i \frac{\left\langle \psi_n^{\lambda} \,\middle|\, \mathbf{p} \,\middle|\, \psi_m^{\lambda} \right\rangle}{\varepsilon_n^{\lambda} - \varepsilon_m^{\lambda}}, \tag{70}$$

arriving at an alternative expression for the derivative of the polarization in a finite crystal:

$$\partial_{\lambda} \mathbf{P}^{\lambda} = \frac{1}{V_c} 2 \operatorname{Im} \sum_{n \leq M} \sum_{m \neq n}^{\infty} \frac{\left\langle \psi_n^{\lambda} \,\middle|\, \mathbf{p} \,\middle|\, \psi_m^{\lambda} \right\rangle \left\langle \psi_m^{\lambda} \,\middle|\, \partial_{\lambda} H^{\lambda} \,\middle|\, \psi_n^{\lambda} \right\rangle}{(\varepsilon_n^{\lambda} - \varepsilon_m^{\lambda})^2}. \tag{71}$$

This expression is also well-defined in an infinite periodic crystal in terms of the Bloch orbitals meaning that the derivative of the polarization represents a pure bulk property, free from the dependence on the termination of the crystal and physics at the surfaces. The expression for

the $\partial_\lambda \mathbf{P}^\lambda$ given as a sum of the matrix elements over occupied and unoccupied states integrated over the Brillouin zone constitutes a key result of the modern theory of electric polarization [8]:

$$\partial_\lambda \mathbf{P}^\lambda = \frac{1}{(2\pi)^3} \sum_{n\leq M} \sum_{m>M}^{\infty} 2\,\mathrm{Im} \int_{\mathrm{BZ}} d\mathbf{k} \frac{\langle \psi^\lambda_{n\mathbf{k}} | \mathbf{p} | \psi^\lambda_{m\mathbf{k}} \rangle \langle \psi^\lambda_{m\mathbf{k}} | \partial_\lambda H^\lambda | \psi^\lambda_{n\mathbf{k}} \rangle}{(\varepsilon^\lambda_{n\mathbf{k}} - \varepsilon^\lambda_{m\mathbf{k}})^2}. \tag{72}$$

Using the latter expression, one can show the Berry phase nature of the electric polarization. Let us rewrite Eq. (72) in terms of the periodic functions $u^\lambda_{n\mathbf{k}}$, rather that $\psi^\lambda_{n\mathbf{k}}$, by noticing that:

$$\left\langle \psi^\lambda_{n\mathbf{k}} \,\middle|\, \mathbf{p} \,\middle|\, \psi^\lambda_{m\mathbf{k}} \right\rangle = \left\langle u^\lambda_{n\mathbf{k}} \,\middle|\, \left[\partial_{\mathbf{k}}, H^\lambda_{\mathbf{k}}\right] \,\middle|\, u^\lambda_{m\mathbf{k}} \right\rangle = \left\langle u^\lambda_{n\mathbf{k}} \,\middle|\, \partial_{\mathbf{k}} H^\lambda_{\mathbf{k}} \,\middle|\, u^\lambda_{m\mathbf{k}} \right\rangle, \tag{73}$$

$$\left\langle \psi^\lambda_{n\mathbf{k}} \,\middle|\, \partial_\lambda H^\lambda \,\middle|\, \psi^\lambda_{m\mathbf{k}} \right\rangle = \left\langle u^\lambda_{n\mathbf{k}} \,\middle|\, \left[\partial_\lambda, H^\lambda(\mathbf{k})\right] \,\middle|\, u^\lambda_{m\mathbf{k}} \right\rangle = \left\langle u^\lambda_{n\mathbf{k}} \,\middle|\, \partial_\lambda H^\lambda_{\mathbf{k}} \,\middle|\, u^\lambda_{m\mathbf{k}} \right\rangle. \tag{74}$$

In turn,

$$\left\langle u^\lambda_{n\mathbf{k}} \right| \partial_{\mathbf{k}} H^\lambda_{\mathbf{k}} \left| u^\lambda_{m\mathbf{k}} \right\rangle = (\varepsilon^\lambda_{n\mathbf{k}} - \varepsilon^\lambda_{m\mathbf{k}}) \left\langle \partial_{\mathbf{k}} u^\lambda_{n\mathbf{k}} \,\middle|\, u^\lambda_{m\mathbf{k}} \right\rangle, \tag{75}$$

$$\left\langle u^\lambda_{n\mathbf{k}} \right| \partial_\lambda H^\lambda_{\mathbf{k}} \left| u^\lambda_{m\mathbf{k}} \right\rangle = (\varepsilon^\lambda_{n\mathbf{k}} - \varepsilon^\lambda_{m\mathbf{k}}) \left\langle \partial_\lambda u^\lambda_{n\mathbf{k}} \,\middle|\, u^\lambda_{m\mathbf{k}} \right\rangle. \tag{76}$$

We can use these relations to reduce the band summation in Eq. (72) to occupied states only [9]:

$$\partial_\lambda \mathbf{P}^\lambda = \frac{1}{(2\pi)^3} 2\,\mathrm{Im} \sum_{n\leq M} \int_{\mathrm{BZ}} d\mathbf{k} \left\langle \partial_{\mathbf{k}} u^\lambda_{n\mathbf{k}} \,\middle|\, \partial_\lambda u^\lambda_{n\mathbf{k}} \right\rangle = -\frac{1}{(2\pi)^3} \sum_{n\leq M} \int_{\mathrm{BZ}} d\mathbf{k}\, \Omega^n_{\lambda\mathbf{k}}. \tag{77}$$

Thus, the derivative of the polarization with respect to λ, Eq. (77), is the Brillouin zone integral of the mixed Berry curvature with respect to $\mathbf{k}$ and λ in the $(\lambda, \mathbf{k})$-space summed over all occupied bands. Since for a non-degenerate band all components of the curvature form are gauge invariant, we conclude that $\partial_\lambda \mathbf{P}^\lambda$ is a gauge invariant quantity.
We can now consider the change in $\mathbf{P}^\lambda$ as λ is varied between, say, 0 and 1, under the condition that at each point along this interval our system stays insulating. This change of the polarization during the λ-evolution of the system can be expressed as:

$$\Delta \mathbf{P} = \int_0^1 \partial_\lambda \mathbf{P}^\lambda d\lambda, \tag{78}$$

where we can substitute now equation (77) and get:

$$\Delta \mathbf{P} = \frac{1}{(2\pi)^3} 2\mathrm{Im} \sum_{n\leq M} \int_0^1 \int_{\mathrm{BZ}} d\mathbf{k} d\lambda \left\langle \partial_{\mathbf{k}} u^\lambda_{n\mathbf{k}} \,\middle|\, \partial_\lambda u^\lambda_{n\mathbf{k}} \right\rangle \tag{79}$$

In the special case of a one-dimensional lattice with a lattice constant a and BZ$= [-\frac{\pi}{a}, \frac{\pi}{a}]$ we rewrite the change in polarization as:

$$\Delta P = \frac{1}{\pi} \mathrm{Im} \sum_{n\leq M} \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \int_0^1 dk\, d\lambda \left\langle \partial_k u^\lambda_{nk} \,\middle|\, \partial_\lambda u^\lambda_{nk} \right\rangle. \tag{80}$$

Topological meaning of electric polarization.
While the geometrical meaning of ΔP as a property related to the Berry curvature in $\mathbf{k}$-space has been clarified, a way of looking at the ΔP as a *topological* property of the occupied states requires to make a compact topological manifold without boundary out of the Brillouin zone. An intuitive way of doing so would be to glue the edges of the BZ which differ by a reciprocal $\mathbf{G}$

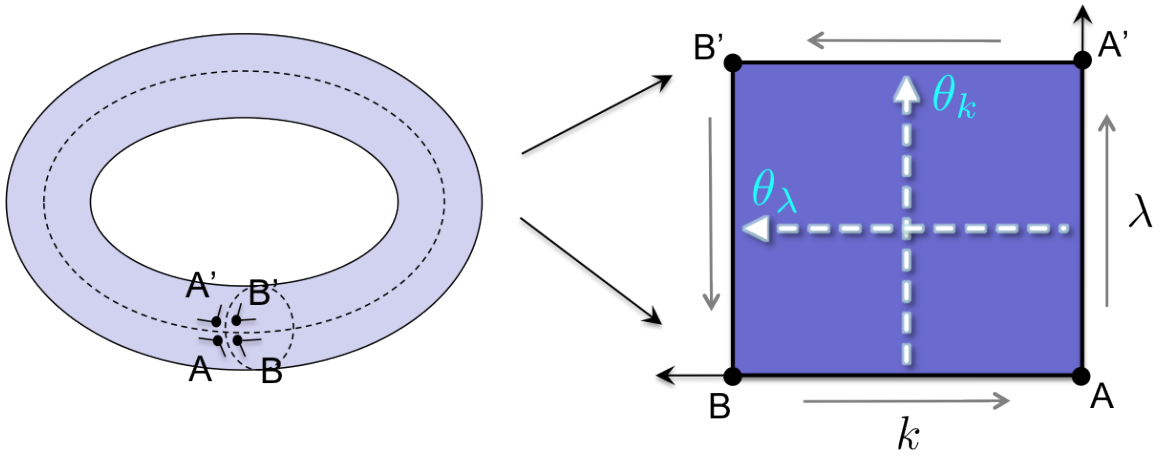


Fig. 2: *Left: torus of states $u^{\lambda}_{n\mathbf{k}}$ where λ runs from 0 to 1 with $H^{\lambda=0}_{\mathbf{k}} = H^{\lambda=1}_{\mathbf{k}}$, and $\mathbf{k}$ runs in the first Brillouin zone. Right: cutting the torus on the left into a square avoiding going through the equators of the torus. This allows to choose smooth and unique gauge everywhere in the square and at its boundaries. For details see text.*

vector, together, i.e., to identify $u_{n\mathbf{k}}$ with $u_{n\mathbf{k}+\mathbf{G}}$. This however cannot be done by just putting $u_{n\mathbf{k}} = u_{n\mathbf{k}+\mathbf{G}}$, since the Hamiltonians at $\mathbf{k}$ and $\mathbf{k}+\mathbf{G}$ are not the same, but are related as follows:

$$H_{\mathbf{k}+\mathbf{G}} = e^{-i\mathbf{G}\mathbf{r}} H_{\mathbf{k}}\, e^{i\mathbf{G}\mathbf{r}}. \tag{81}$$

From the latter relation we can conclude that the set of instantaneous solutions of $H_{\mathbf{k}+\mathbf{G}}$ can be always chosen to be as follows:

$$\{u_{n\mathbf{k}+\mathbf{G}}\} = e^{-i\mathbf{G}\mathbf{r}}\{e^{i\theta_n} u_{n\mathbf{k}}\}, \tag{82}$$

where θ_n are band-dependent constants which stand for the obvious $\mathbb{U}(1)$ gauge freedom which we considered previously. It can be shown straightforwardly that the Berry phase γ_n computed along a path which connects the $\mathbf{k}$ and $\mathbf{k}+\mathbf{G}$ points is a gauge-invariant quantity, without the path of integration being formally closed, and the whole geometrical machinery we presented in section 2 can be also applied here. Indeed, imagine two families of Hamiltonians, $H(\lambda)$ and $H'(\lambda)$, which differ by a *constant* λ-independent unitary transformation:

$$H'(\lambda) = UH(\lambda)U^{\dagger}, \tag{83}$$

with U and $H(\lambda)$ not necessarily commuting. Then it is clear that if at a point λ the set of instantaneous solutions of $H(\lambda)$ is $\{|n\lambda\rangle\}$, then the set $\{|n\lambda\rangle'\} = \{U\,|n\lambda\rangle\}$ presents a set of instantaneous solutions of $H'(\lambda)$. If in the vicinity of certain λ function $|n\lambda\rangle$ is smooth, so will be $|n\lambda\rangle'$. In terms of a gauge transformation, the two sets are connected by a trivial, λ-independent gauge transformation. The connections of H and H' can be also compared:

$$\mathcal{A}'^{n} = i\,\langle n\lambda'|\partial_\lambda|n\lambda'\rangle\,d\lambda = i\,\langle Un\lambda|\partial_\lambda|Un\lambda\rangle\,d\lambda = i\,\langle n\lambda|\partial_\lambda|n\lambda\rangle\,d\lambda = \mathcal{A}^{n}. \tag{84}$$

This means that the Berry phase of any closed path $\mathcal{C}$ for both Hamiltonians is the same. This is actually true even for any path $\mathcal{C}$ which is not closed. In case of Bloch electrons, the role of U is played by $e^{-i\mathbf{G}\mathbf{r}}$.

The most convenient way to deal with the topological properties of the BZ is to consider what we call the *folded BZ*. At each point $\mathbf{k}$ in the first BZ we consider the *equivalence classes* of lattice periodic functions which differ from each other by $e^{i\mathbf{G}\mathbf{r}}$, where $\mathbf{G}$ is a reciprocal lattice vector. The equivalence class which corresponds to a certain $u_{n\mathbf{k}}$ is given by:

$$\hat{u}_{n\mathbf{k}} = \{e^{-i\mathbf{G}\mathbf{r}} u_{n\mathbf{k}},\ \forall \mathbf{G}\}. \tag{85}$$

Physically, $\hat{u}_{n\mathbf{k}}$ is an object which accumulates all instantaneous solutions at points $\mathbf{k}+\mathbf{G}$ in the reciprocal space and folds them into the first BZ. $u_{n\mathbf{k}}$ is called the *representative* of this class. Respectively, we can consider the equivalence classes of the Hamiltonians,

$$\hat{H}_{\mathbf{k}} = \{e^{-i\mathbf{G}\mathbf{r}} H_{\mathbf{k}}\, e^{i\mathbf{G}\mathbf{r}},\ \forall \mathbf{G}\}, \tag{86}$$

for which corresponding $\hat{u}_{n\mathbf{k}}$ are the instantaneous solutions:

$$\hat{H}_{\mathbf{k}} \hat{u}_{n\mathbf{k}} = \varepsilon_{n\mathbf{k}} \hat{u}_{n\mathbf{k}}. \tag{87}$$

Importantly, as opposed to $H_{\mathbf{k}}$, the corresponding equivalence class $\hat{H}_{\mathbf{k}}$ is periodic in reciprocal space, since $\hat{H}_{\mathbf{k}+\mathbf{G}} = \hat{H}_{\mathbf{k}}$. It is easy to show that the Berry phase theory from above can be transparently formulated in terms of $\hat{u}_{n\mathbf{k}}$'s rather then $u_{n\mathbf{k}}$'s. If the gauge freedom in the choice of $\hat{u}_{n\mathbf{k}}$'s is introduced as

$$\hat{u}'_{n\mathbf{k}} = e^{i\theta_n(\mathbf{k})} \hat{u}_{n\mathbf{k}} := \{e^{i\theta_n(\mathbf{k})} e^{-i\mathbf{G}\mathbf{r}} u_{n\mathbf{k}},\ \forall \mathbf{G}\}, \tag{88}$$

then we can show, using (84), that the Berry connection $\hat{\mathcal{A}}^n$ can be defined uniquely for $\hat{u}_{n\mathbf{k}}$ as

$$\hat{\mathcal{A}}^n = i \langle \hat{u}_{n\mathbf{k}} | \partial_{\mathbf{k}} | \hat{u}_{n\mathbf{k}} \rangle\, d\mathbf{k} := i \langle u_{n\mathbf{k}} | \partial_{\mathbf{k}} | u_{n\mathbf{k}} \rangle\, d\mathbf{k}, \tag{89}$$

with gauge transformation reading in analogy to (16):

$$\hat{\mathcal{A}}'^n = \hat{\mathcal{A}}^n - d\theta_n. \tag{90}$$

Finally, in our construction of the folded BZ we identify the BZ zone boundaries which differ from each other by any $\mathbf{G}$, which can be done since $\hat{H}_{\mathbf{k}}$ is a periodic function in reciprocal space. In case of a 2D BZ this looks like a torus, see Fig. 2, and results in a topological construction of a closed compact manifold. As we remember from the example of spin-$\frac{1}{2}$ in magnetic field the Chern number was ultimately related to our inability of constructing a global smooth and unique gauge on S^2. The same is true for our folded BZ, as we shall see below. If, starting from a point $\mathbf{k}$ in a folded BZ, we could go around our torus along one of the diameters smoothly and uniquely on the whole closed path, obviously, we would arrive at the condition that:

$$\hat{u}_{n\mathbf{k}} \longrightarrow \hat{u}_{n\mathbf{k}} \quad \Leftrightarrow \quad u_{n\mathbf{k}+\mathbf{G}} = e^{-i\mathbf{G}\mathbf{r}} u_{n\mathbf{k}},\ \forall \mathbf{G}. \tag{91}$$

The choice of such a gauge in the whole folded BZ, if it is possible to make, is called the *periodic gauge*. If the Chern number of our manifold is non-zero, such a choice is impossible to make, and, generally speaking:

$$\hat{u}_{n\mathbf{k}} \longrightarrow e^{i\theta_n} \hat{u}_{n\mathbf{k}}, \tag{92}$$

where (generally $\mathbf{k}$-dependent) θ_n is not necessarily a multiple of 2π. Next, lets analyze in detail the relation between the Chern number and the phase θ_n for the case of electric polarization (80).

Chern number and the 2-point formula.
Let us consider the case of $H(\lambda = 0) = H(\lambda = 1)$. For the (λ, k) situation of Eq. (80) the general relation between the instantaneous solutions at the opposite sides of the square read, analogously to (92):

$$u_k^{\lambda=1} = e^{i\theta_k} u_k^{\lambda=0}, \tag{93}$$

$$u_{\frac{\pi}{a}}^{\lambda} = e^{-iGx}\, e^{i\theta_\lambda} u_{-\frac{\pi}{a}}^{\lambda}, \tag{94}$$

where we skip for simplicity the band index n. Let us now try to evaluate the ΔP. Following Fig. 2 we can cut the torus such that on the boundaries of the square $AA'B'B$ we have a smooth choice of the u_k^λ. This is equivalent to the reasonable assumption that all the "non-smoothness" has been restricted to the infinitesimally small "belts" around the two equators of the torus, which are analogous to an infinitesimally small "belt" around the equator of the S^2 in case of spin-$\frac{1}{2}$ in magnetic field. Then the integral of the mixed Berry curvature over the torus, given by Eq. (80), can be rewritten as an integral of the Berry connection along the path $\mathcal{C} = AA'B'B$, Fig. 2:

$$\Delta P = \frac{1}{2\pi}\left(\int_0^1 d\lambda \left[\mathcal{A}_\lambda(\lambda, -\frac{\pi}{a}) - \mathcal{A}_\lambda(\lambda, \frac{\pi}{a})\right] + \int_{-\pi/a}^{\pi/a} dk \left[\mathcal{A}_k(1,k) - \mathcal{A}_k(0,k)\right]\right), \tag{95}$$

where $\mathcal{A}_\lambda = i\left\langle u_{n\mathbf{k}}^\lambda \,\middle|\, \partial_\lambda u_{n\mathbf{k}}^\lambda \right\rangle$ and $\mathcal{A}_k = i\left\langle u_{n\mathbf{k}}^\lambda \,\middle|\, \partial_{\mathbf{k}} u_{n\mathbf{k}}^\lambda \right\rangle$ are the components of the Berry connection. Using Eqs. (93) and (94), it is easy to see that the expression for ΔP is reduced to:

$$\Delta P = \frac{1}{2\pi}\left(\int_0^1 \frac{\partial\theta_\lambda}{\partial\lambda} d\lambda - \int_{-\pi/a}^{\pi/a} \frac{\partial\theta_k}{\partial k} dk\right). \tag{96}$$

Since the functions θ are smooth along the boundary, we come to the conclusion that

$$\Delta P = \frac{\theta_\lambda(1) - \theta_\lambda(0) - \theta_k(\pi/a) + \theta_k(-\pi/a)}{2\pi}. \tag{97}$$

On the other hand, according to Eqs. (93) and (94), upon the $A \to A' \to B' \to B$ evolution the wavefunction $u_{-\pi/a}^{\lambda=0}$ is returned into the wavefunction $e^{i(\theta_\lambda(1)-\theta_\lambda(0)-\theta_k(\pi/a)+\theta_k(-\pi/a))} u_{-\pi/a}^{\lambda=0}$. Since we chose our gauge smoothly along this path,

$$u_{-\pi/a}^{\lambda=0} = e^{i(\theta_\lambda(1)-\theta_\lambda(0)-\theta_k(\pi/a)+\theta_k(-\pi/a))} u_{-\pi/a}^{\lambda=0}, \tag{98}$$

meaning that $\theta_\lambda(1) - \theta_\lambda(0) - \theta_k(\pi/a) + \theta_k(-\pi/a)$ has to be a multiple of 2π, and ΔP has to be an integer. In fact, it is the first Chern number of the system, as we have seen in the previous section:

$$\Delta P = C \quad \longrightarrow \quad \Delta P = \sum_n C_n, \tag{99}$$

where C_n is the first Chern number of band n in the (λ, k) space and the arguments above present the proof of the quantization of the Chern number for this particular situation. This proof is however more general than the one we provided for the spin-$\frac{1}{2}$ in magnetic field, since it does not rely on the precise expressions for the Berry connection (49). Depending on the Hamiltonian of the system, H_k^λ, the Chern number of the system, which *uniquely* in mathematical sense depends on the system, can be either zero, or it can be non-zero. In the latter case we say that

we encounter a situation of a *Chern insulator*, as opposed to the trivial insulator with the Chern number zero. In the previous section we provided an example of a Hamiltonian which is a Chern insulator, namely, a spin-$\frac{1}{2}$ particle in magnetic field, in which case the sphere S^2 played the role of the (λ, k) space. In case when the system is not a Chern insulator the smooth and unique choice of the wavefunction can be found everywhere on the torus, which corresponds to the case of $\theta_k \equiv \theta_\lambda \equiv 0$ and $u_k^{\lambda=1} = u_k^{\lambda=0}, u_{\pi/a}^{\lambda} = e^{-iGx} u_{-\pi/a}^{\lambda}$, as discussed above. In the next section we elaborate in more detail on Chern insulators in two-dimensional reciprocal space.
Let us assume that we can choose the periodic gauge in the reciprocal space for each λ. In this case $\theta_\lambda \equiv 0$ and the change of the polarization is given by:

$$\Delta P = \frac{\theta_k(-\pi/a) - \theta_k(\pi/a)}{2\pi} = -\int_{-\pi/a}^{\pi/a} \frac{\partial \theta_k}{\partial k} dk. \tag{100}$$

We can even go further back to write the so-called *2-point formula* for the change of polarization:

$$\Delta P = P_1 - P_0 = \frac{i}{2\pi} \int_{-\pi/a}^{\pi/a} dk \, \left\langle u_k^{\lambda} \,\middle|\, \partial_k u_k^{\lambda} \right\rangle^{\lambda=1} - \frac{i}{2\pi} \int_{-\pi/a}^{\pi/a} dk \, \left\langle u_k^{\lambda} \,\middle|\, \partial_k u_k^{\lambda} \right\rangle^{\lambda=0}. \tag{101}$$

The latter equation absolutely correctly gives the value of ΔP, if we remember that we started from Eq. (100). What is meant is that very often, especially in practical calculations of the polarization, the expresion (101) is considered separately without its reference to Eq. (100). This means that the sets of $\{u_{nk}^{\lambda=0}\}$ and $\{u_{nk}^{\lambda=1}\}$ wavefunctions are calculated separately omitting the λ-evolution, and the value of ΔP is then calculated as in (101). In this way we do not built explicitly a unique smooth function θ_k and it can in principal take any values. Applying expression (101) properly would mean, given a certain set of u_{nk} in the BZ for $\lambda = 0$, to go smoothly by hand from $u_{nk}^{\lambda=0}$ across u_{nk}^{λ} to construct $u_{nk}^{\lambda=1}$, and to use such constructed $u_{nk}^{\lambda=1}$ in (101). If we do not do this, then the only thing we know is that due to periodic boundary conditions $e^{i\theta_k(\pi/a)} = e^{i\theta_k(-\pi/a)}$, meaning that the difference of $\theta_k(-\pi/a) - \theta_k(\pi/a)$ is defined up to a multiple of 2π, with ΔP as an integer, but undetermined in value. The quantity (general to multi-dimensional λ and $\mathbf{k}$)

$$\mathbf{P}^{\lambda} = \frac{i}{(2\pi)^3} \sum_{n \le M} \int_{\mathrm{BZ}} d\mathbf{k} \left\langle u_{n\mathbf{k}}^{\lambda} \,\middle|\, \partial_{\mathbf{k}} u_{n\mathbf{k}}^{\lambda} \right\rangle \tag{102}$$

is called the *electric polarization* and the two-point formula for $\Delta\mathbf{P}$ can be generalized in exact analogy to Eq. (101) to higher dimensions. It is straightforward to show that if $H(\lambda = 0) \neq H(\lambda = 1)$, Eq. (95) is still valid. This means, that although θ_k cannot be appropriately defined anymore, given the possibility of a periodic gauge in $\mathbf{k}$-space for each λ, and a smooth connection between $u_{nk}^{\lambda=0}$ and $u_{nk}^{\lambda=1}$, the two-point formula can be also applied, although the value of ΔP is not anymore quantized.

Adiabatic pumping and velocity of Bloch electrons.
Evaluation of the velocity of electronic states requires going beyond the adiabatic approximation for the evoluted wavefunction (11-12). Up to first order in transition frequencies (10) equations (9) can be solved to yield:

$$|\psi(t)\rangle = e^{-i\int_0^t \varepsilon_n(\tau)d\tau} \left[|n\lambda(t)\rangle - i \sum_{m \neq n} \frac{\langle m\lambda(t)|\partial_t|n\lambda(t)\rangle}{\varepsilon_n(t) - \varepsilon_m(t)} |m\lambda(t)\rangle \right]. \tag{103}$$

The latter expression goes beyond the adiabatic approximation in that it also includes the "smearing" of the wavefunction which was initially in state n, over other eigenstates of the Hamiltonian during the time-propagation. In deriving (103) the parallel transport gauge (19) was used, which locally eliminates the geometric phase contribution to the evolution of ψ. This is fine, however, since the final expression will be locally gauge-invariant. Let us now try to evaluate the velocity of electronic states in a crystal during time-evolution given by change in parameter λ, which could be time itself. Namely, at time $t = 0$ at a certain $\mathbf{k}$ we start with a wavefunction $\psi_{n\mathbf{k}}^{\lambda(t=0)}$ which is analogous to state $|n\lambda(0)\rangle$ from above, and we evaluate the velocity of the evoluted wavefunction at infinitesimally small time $t > 0$. It is most convenient to assume at the moment that λ enters the Hamiltonian according to (66) via the crystal potential, in which case $\mathbf{k}$ stays constant as λ is changed. To distinguish the properly evoluted states from the instantaneous states $\psi_{n\mathbf{k}}^{\lambda(t)}$ and $u_{n\mathbf{k}}^{\lambda(t)}$ (playing the role of $|n\lambda(t)\rangle$ above), we denote them by $\tilde{\psi}_{n\mathbf{k}}^{\lambda}$ and $\tilde{u}_{n\mathbf{k}}^{\lambda}$. Then the average velocity of state $\tilde{\psi}_{n\mathbf{k}}^{\lambda}$ at point $\lambda = \lambda(t)$ is given by (using also (73)):

$$\mathbf{v}_{n\mathbf{k}}^{\lambda} = \left\langle \tilde{\psi}_{n\mathbf{k}}^{\lambda} \middle| i\left[H^{\lambda}, \mathbf{r}\right] \middle| \tilde{\psi}_{n\mathbf{k}}^{\lambda} \right\rangle = \left\langle \tilde{u}_{n\mathbf{k}}^{\lambda} \middle| \partial_{\mathbf{k}} H_{\mathbf{k}}^{\lambda} \middle| \tilde{u}_{n\mathbf{k}}^{\lambda} \right\rangle. \tag{104}$$

On the other hand from (103) it follows that:

$$|\tilde{u}_{n\mathbf{k}}^{\lambda}\rangle = e^{-i\int_0^t \varepsilon_{n\mathbf{k}}^{\lambda(\tau)}\, d\tau} \left[|u_{n\mathbf{k}}^{\lambda}\rangle - i \sum_{m\neq n} |u_{m\mathbf{k}}^{\lambda}\rangle \frac{\left\langle u_{m\mathbf{k}}^{\lambda} \middle| \partial_{\lambda} u_{n\mathbf{k}}^{\lambda} \right\rangle}{\varepsilon_{n\mathbf{k}}^{\lambda} - \varepsilon_{m\mathbf{k}}^{\lambda}} \dot{\lambda} \right], \tag{105}$$

where $\dot{\lambda} = \partial_t \lambda(t)$. Then,

$$\mathbf{v}_{n\mathbf{k}}^{\lambda} = \partial_{\mathbf{k}} \varepsilon_{n\mathbf{k}}^{\lambda} + 2\mathrm{Im} \sum_{m\neq n} \left\langle u_{n\mathbf{k}}^{\lambda} \right| \partial_{\mathbf{k}} H_{\mathbf{k}}^{\lambda} \left| u_{m\mathbf{k}}^{\lambda} \right\rangle \frac{\left\langle u_{m\mathbf{k}}^{\lambda} \middle| \partial_{\lambda} u_{n\mathbf{k}}^{\lambda} \right\rangle}{\varepsilon_{n\mathbf{k}}^{\lambda} - \varepsilon_{m\mathbf{k}}^{\lambda}} \dot{\lambda}. \tag{106}$$

Using Eqs. (75) from before, we arrive at the following two equivalent expressions for the velocity of the state

$$\mathbf{v}_{n\mathbf{k}}^{\lambda} = \partial_{\mathbf{k}} \varepsilon_{n\mathbf{k}}^{\lambda} + 2\mathrm{Im} \sum_{m\neq n} \frac{\left\langle u_{n\mathbf{k}}^{\lambda} \middle| \partial_{\mathbf{k}} H_{\mathbf{k}}^{\lambda} \middle| u_{m\mathbf{k}}^{\lambda} \right\rangle \left\langle u_{m\mathbf{k}}^{\lambda} \middle| \partial_{\lambda} H_{\mathbf{k}}^{\lambda} \middle| u_{n\mathbf{k}}^{\lambda} \right\rangle}{(\varepsilon_{n\mathbf{k}}^{\lambda} - \varepsilon_{m\mathbf{k}}^{\lambda})^2} \dot{\lambda}, \tag{107}$$

or, according to (24),

$$\mathbf{v}_{n\mathbf{k}}^{\lambda} = \partial_{\mathbf{k}} \varepsilon_{n\mathbf{k}}^{\lambda} - \Omega_{\lambda\mathbf{k}}^{n} \dot{\lambda}. \tag{108}$$

In this expression the first term of the right hand side is the *group velocity*, which is present also when λ is constant and the stationary state $\psi_{n\mathbf{k}}^{\lambda}$ evolves in time according to (5). The change in λ gives, on the other hand, rise to the so-called *anomalous velocity*, which is expressed in terms of the mixed Berry curvature $\Omega_{\lambda\mathbf{k}}^{n}$. Clearly, the anomalous velocity of a state depends on how fast the parameter λ is changed in time. We can now evaluate the current density due to all occupied states changing of λ in time induces. It is given by:

$$\mathbf{J}^{\lambda} = -\frac{1}{(2\pi)^d} \dot{\lambda} \sum_{n\leq M} \int_{\mathrm{BZ}} \Omega_{\lambda\mathbf{k}}^{n}\, d\mathbf{k}, \tag{109}$$

while the contribution due to the group velocity of the states clearly vanishes, since the band energies are periodic functions of $\mathbf{k}$. The polarization charge pumped during the evolution from

λ_1 to λ_2 is given by the time integral of the current density from above from t_1 to t_2 with $\lambda(t_1) = \lambda_1$ and $\lambda(t_2) = \lambda_2$, which amounts to:

$$\Delta \mathbf{P} = \int_{t_1}^{t_2} \mathbf{J}^{\lambda(t)}\, dt = -\frac{1}{(2\pi)^d} \sum_{n \leq M} \int_{\lambda_1}^{\lambda_2} \int_{\mathrm{BZ}} \Omega^n_{\lambda \mathbf{k}}\, d\mathbf{k} d\lambda, \tag{110}$$

which gives us exactly the change in the polarization as given by Eq. (79). This expression presents a theoretical justification for the interpretation of electric polarization in terms of experimentally measured charge current. In the special case of one dimension, as we have seen previously, the transported through the system charge during cyclic adiabatic evolution is thus quantized, which leads to the phenomenon of *adiabatic pumping*.

Symmetry properties of the Berry curvature.
Based on general symmetry arguments, it is straightforward to show that the Berry curvature in $\mathbf{k}$-space obeys the following symmetry properties: (i) in the presence of time-reversal symmetry, $\Omega^n(-\mathbf{k}) = -\Omega^n(\mathbf{k})$, while (ii) in the presence of the space inversion symmetry, $\Omega^n(-\mathbf{k}) = \Omega^n(\mathbf{k})$. This means that when both space and time inversion symmetry are present in a solid, the Berry curvature at each $\mathbf{k}$ is identically zero. In case of materials, which display non-zero electric polarization, the Berry curvature is non-trivial owing to the breaking of inversion symmetry. On the other hand, in materials which exhibit spontaneous magnetization, such as ferromagnets and antiferromagnets, the non-trivial Berry curvature is due to breaking of time-inversion symmetry. We will focus on the latter case in the next sections.

3.3 Chern insulators and (quantum) anomalous Hall effect

In the remainder of these notes we will call the *Chern insulator* a two-dimensional (2D) insulating solid with Hamitonian $H(\mathbf{k})$ whose first Chern number, determined as an integral of the $\mathbf{k}$-space Berry curvature over the Brillouin zone, is an integer *non-zero* number. The proof of the fact that the Chern number is integer for a two-dimensional insulator we have provided in the previous section, where one has to replace k with k_x and λ with k_y. The condition of the periodicity of the Hamiltonian with respect to λ which we used to glue the square in Fig. 2 into a torus can be satisfied in reciprocal space in k_y-direction following the folding procedure we decribed above, which allows us to look at the 2D Brillouin zone as a torus. The condition that the Chern number is non-zero means that the topological properties of our system are non-trivial, and that the wavefunction $u_{n\mathbf{k}}$ acquires a "twist" as we cross at least one of the equators, which manifests in the non-zeroness of θ_{k_x} or θ_{k_y}. Chern insulators present an example of a system for which the periodic gauge in both k-directions cannot be found.
One of the most remarkable properties of Chern insulators is the quantization of their transverse Hall conductance. From the general linear response formalism as well as from the semiclassical electron dynamics which we consider later, it follows that Hall conductance of a 2D system is given by:

$$\sigma_{xy} = \frac{1}{2\pi} \int_{\mathrm{BZ}} d\mathbf{k}\, \Omega_{xy}(\mathbf{k}) = \frac{1}{2\pi} \int_{\mathrm{BZ}} d\mathbf{k}\, \Omega_k = \sum_n C_n, \tag{111}$$

where C_n is the first Chern number of the occupied band n. From the point of view of adiabatic pumping we considered previously, the quantization of the Hall conductance is due to the quantized charge which is pumped through the one-dimensional system along x described by

Hamiltonian $H(k_x, k_y)$ as the parameter k_y is varied from one side of the BZ to the other. This corresponds to the situation of varying λ between 0 and 1 in Eq. (80).
Let us take a look at the mechanisms which can lead to appearance of Chern insulators. For this purpose we consider first the case of only two bands in $\mathbf{k}$-space (keeping in mind for future reference that the role of $\mathbf{k}$ can be replaced by any parameter). Neglecting the constant term, which does not change the topological properties and just shifts the energy, a generic 2D Hamiltonian reads:

$$H(\mathbf{k}) = \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}, \tag{112}$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices. At each point in $\mathbf{k}$ the energetic spectum is given by two eigenvalues ε_+ and ε_- which are given by $\pm|\mathbf{d}(\mathbf{k})| = \pm d(\mathbf{k})$. Since the Berry curvature summed over both bands is always zero, we will consider only lowest of the bands, ε_-. We also suppose that both bands are separated from each other by a gap. By comparing the Hamiltonian (112) to that of spin-$\frac{1}{2}$ in magnetic field, it is clear that the physics of the problem is governed by the Dirac monopole. Indeed, the Berry phase which is accumulated when going along a closed loop $\mathcal{C}$ in $\mathbf{k}$-space can be easily related to that picked up when going along a loop $\chi(\mathcal{C})$ on a unit sphere S^2, where χ maps $\mathbf{k}$ to the point $\mathbf{d}/d$ on S^2, see Fig. 1:

$$\chi : \mathbf{k} \longrightarrow \mathbf{n} = \mathbf{d}/d, \quad \mathbf{n} \in S^2. \tag{113}$$

The Berry curvature of the problem on S^2 is given by a field of the Dirac monopole with quantized charge at the origin, where the two bands touch each other and $\varepsilon_+ = \varepsilon_-$. As we remember from the considerations of S^2, for a given spin we cannot choose the smooth gauge of our wavefunctions in the BZ and we have to "glue" them together at the diameter of the sphere. The Berry curvature in the $\mathbf{k}$-space is readily obtained from the Berry curvature of the Dirac monopole multiplied with the Jacobian of χ:

$$\Omega_{xy}(\mathbf{k}) = -\,\mathbf{n} \cdot \left(\partial_{k_x}\mathbf{n} \times \partial_{k_y}\mathbf{n}\right)/2 = \boldsymbol{\Omega}^{\sigma}_{\mathbf{n}} \cdot \left(\partial_{k_x}\mathbf{d} \times \partial_{k_y}\mathbf{d}\right), \tag{114}$$

where $\boldsymbol{\Omega}^{\sigma}_{\mathbf{n}}$ is that given by Eq. (49) for $\sigma = -1$. As we learned from mathematics, the first Chern number obtained as an integral of the Berry curvature over a *compact* manifold such as a 2D BZ, has to be an integer. For (114) it is called the *winding number* and it stands for the number of times that the field $\mathbf{d}(\mathbf{k})$ winds around the S^2 as $\mathbf{k}$ is varied, as can be intuitively understood and computed explicitly.
A good example when the integer quantization is violated is the massive Dirac Hamiltonian, for which $d_x = k_x, d_y = k_y$ and $d_z = m$:

$$H(\mathbf{k}) = k_x\sigma_x + k_y\sigma_y + m\sigma_z. \tag{115}$$

This Hamiltonian both with zero and non-zero m is of great importance for studying topological properties of solids. The Berry curvature can be evaluated analytically:

$$\Omega_{xy}(\mathbf{k}) = m/[2(m^2 + k^2)^{3/2}], \tag{116}$$

which leads to the following Hall conductance when integrated over the whole infinite BZ: $\sigma_{xy} = \mathrm{sgn}(m)/2$, leading thus to half-integer-quantized values. This seems to be in contradiction with our expectation of integer quantization. The reason for this is that we have here a situation of a non-compact manifold for $\mathbf{k}$-space, namely, $\mathbb{R}^2$. By looking at the distribution of $\mathbf{d}$ over $\mathbb{R}^2$ we realize that for a given m the sign of d_z remains constant and the vector $\mathbf{d}$

spans only half of the S^2, leading to the so-called meron situation. This results in half-integer conductance. In a realistic situation of a lattice which provides a periodic lattice potential and finite band width, the BZ is compact, and the sign of d_z changes at other points of the BZ, where the bands bend down in order to assure a finite band width. Such regions provide the other half of the conductance, vector $\mathbf{d}$ spans the other half of the sphere and the resulting conductance indeed becomes integer. The moral of this story is that in order to determine the Chern number of a material, it is not enough to concentrate only on the Dirac-like localized points in the BZ which can provide very large contribution to the Berry curvature, and the whole band structure of the crystal has to be considered.

At the point $m = 0$ in the Hamiltonian (115) there is a point of degeneracy in the spectrum at $k_x = k_y = 0$ and the transition between two Chern insulator phases for $m < 0$ and $m > 0$ occurs. Such situation is rather typical and can be generalized to the case of a 2D+1 Hamiltonian $H(k_x, k_y, \lambda) = H(\boldsymbol{\varkappa})$, where $\boldsymbol{\varkappa} = (k_x, k_y, \lambda)$ and λ is a certain parameter, such as m in the Dirac model, values of hoppings in the lattice Hamiltonian, value of the spin-orbit strength, exchange field, etc. According to the theorem by Bellisard [10], the change of the (k_x, k_y) Chern number when going through a point of degeneracy at a certain value of $\boldsymbol{\varkappa}^*$ (such as $(0, 0, 0)$ in the Dirac model) is given by the so-called *Berry index*:

$$\mathrm{Indx_B} = \frac{1}{2\pi} \int_{S^2} \Omega_k(\boldsymbol{\varkappa}) d\boldsymbol{\varkappa}, \tag{117}$$

where S^2 is an infinitesimally small sphere which encloses $\boldsymbol{\varkappa}^*$. Interestingly, the Berry index determines the change in the 2D Hall conductance irrespective of the compactness of the $\mathbf{k}$-space. E.g. for the massive Dirac Hamitonian (115) the Berry index can be evaluated to be $+1$ at the point $(k_x = 0, k_y = 0, m = 0)$. Moreover, it can be shown that the value of the Berry index is pre-determined by the band dispersion in the vicinity of $\boldsymbol{\varkappa}^*$ [11]. Infinitesimally closely to $\boldsymbol{\varkappa}^*$ we can approximate our two-band Hamiltonian as (omitting again the constant energy term):

$$H(\boldsymbol{\varkappa}) = h(\boldsymbol{\varkappa} - \boldsymbol{\varkappa}^*) \cdot \boldsymbol{\sigma}. \tag{118}$$

If the dispersion of $h(\boldsymbol{\varkappa} - \boldsymbol{\varkappa}^*)$ is linear, then the Berry index assumes the values of ± 1, while if it is quadratic, $\mathrm{Indx_B}$ is ± 2. For the massive Dirac model (115) Hamiltonian $h(\boldsymbol{\varkappa} - \boldsymbol{\varkappa}^*)$ is linear in $\boldsymbol{\varkappa} - \boldsymbol{\varkappa}^*$ which results in the change of Hall conductance by $+1$ as the mass m changes sign. It is probably worthy to note here that when the role of the parameter λ is played by the k_z Bloch vector of a 3D Hamiltonian $H(k_x, k_y, k_z)$, such points of degeneracy $\boldsymbol{\varkappa}^* = \mathbf{k}^*$ will be called here *Weyl points*. If such points happen to be present at the Fermi energy of a material with no other bands crossing it, such material is called a *topological metal*, or *Weyl semimetal*. The physics of topological metals is an exciting emerging field of topological solid state physics.

Historically, Thouless and co-workers [12] were the first to demonstrate that Chern insulators can arise for periodic 2D solids exposed to an external magnetic field. Obtained in such a way Chern insulator can be named the *quantum Hall insulator*, since the quantization of Hall conductance in such Chern insulators was observed in measurements of the integer quantum Hall effect. The quantum Hall insulators are to be distinguished from spontaneous Chern insulators, for which the Chern insulator state is realized without external fields, and the breaking of time-reversal symmetry (necessary for non-zero Berry curvature) is the intrinsic property of the material due to e.g. formation of local spin moments. We will refer to spontaneous Chern insulators as *quantum anomalous Hall insulators* (QAH insulators), since the quantization of conductance in QAH insulators is observed by measuring the anomalous Hall effect (AHE) for

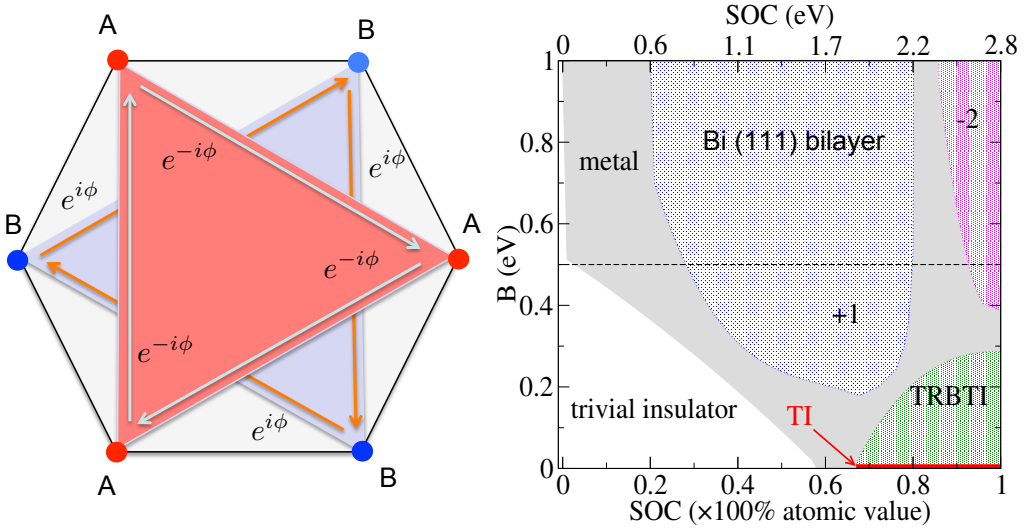


Fig. 3: *Left: Real-space schematic representation of the Haldane model. Right: The phase diagram of the Bi(111) bilayer with respect to the strength of atomic SOC and magnitude of exchange field B. Numbers denote the Chern number in the quantum anomalous Hall phase, TI stands for the topological insulator phase, while TRBTI stands for the time-reversal broken topological insulator phase. For more details see text and Ref. [15].*

which external magnetic field plays a secondary role [13]. The corresponding effect is called the *QAH effect*.

Although the non-zero Chern number is due to non-trivial distribution of the Berry curvature in $\mathbf{k}$-space, a fruitful analysis of QAH phases can be achieved in real space by considering various mechanisms of electron hopping and interactions on a lattice, and corresponding tight-binding Hamiltonians. The first lattice model for a QAH insulator was given by Haldane [14]. The Hamiltonian on the honeycomb lattice within the spinless *Haldane model* looks very simple:

$$H = t_1 \sum_{\langle i,j \rangle} c_i^\dagger c_j + t_2 \sum_{\langle\langle i,j \rangle\rangle} e^{-i\sigma\phi} c_i^\dagger c_j, \tag{119}$$

where the first term corresponds to the hopping between the nearest neighbors, while the second term corresponds to the hopping between the next-nearest neighbors. The key feature of the Haldane model is that the next-nearest neighbor hopping t_2 acquires a complex phase $e^{-i\sigma\phi}$, where $\sigma = +1$ for the hopping on the A-sublattice, while $\sigma = -1$ for the hopping within the B-sublattice. The effect of aquiring a complex phase during electron hopping can be seen as a result of a fictitious magnetic field with the vector potential $\mathbf{A}(\mathbf{r})$: $e^{-i\phi} = e^{-i\int d\mathbf{r}\mathbf{A}(\mathbf{r})}$, where the integral in taken along the shortest path which connect the next-nearest neighbor sites. As the electron completes a closed path when hopping on the corresponding sublattice (see triangles in Fig. 3), it accumulates a phase which is proportional to the flux of the magnetic field through the corresponding triangle, in analogy to the AB-effect we considered before.[2] Since this phase is opposite for electrons of two sublattices, the total field acting on electrons averages to zero within the unit cell.[3] Haldane showed by Fourier transforming the lattice Hamiltonian to the $\mathbf{k}$-

[2]The AB-effect with magnetic field opposite for electrons of different spin, and not sublattice, will re-appear again in the context of the topological Hall effect.

[3]That is why the Haldane model is often called a model for a quantum Hall effect with zero magnetic field.

space that the Chern number of this model equals $+1$ for $-\pi < \phi < 0$, and -1 for $0 < \phi < \pi$. The point $\phi = 0$ thus gives rise to Weyl points in $(\mathbf{k}, \phi)$-space. Conceptually, the suggestion of his model by Haldane in 1988 stands at the origin of the tremendous advances in topological condensed matter physics which followed.

One of the reasons for this is that the mechanism which gives non-zero Chern number within the Haldane model can be realized in actual materials, with intrinsic spin-orbit interaction (SOI) (discussed in detail by Gustav Bihlmayer in manuscript A10 of this book) playing the role of the source of "fictitious" magnetic field which provides non-trivial band topology. To briefly demonstrate how this comes about we focus here on one of the many possible examples considered by now in the literature [5]. Namely, we will consider the p_z orbitals on a strongly buckled honeycomb lattice of space-inversion symmetric (111) bilayer of Bismuth, see Fig. 4. The nearest-neighbor tight-binding multi-orbital lattice Hamiltonian of this system reads:

$$H \;=\; \sum_{ij} t_{ij} c_i^\dagger c_j + \sum_i c_i^\dagger (\varepsilon_i \mathbb{I} + B\sigma_z) c_i \;+ H_{\rm SOC}, \tag{120}$$

where the first term is the kinetic nearest-neighbor hopping between generally different multiple s and p (and d or f in transition and rare-earth metals) orbitals. The second term stands for an orbital on-site energy ε_i and the interaction with the Zeeman exchange field B directed along the z-axis, with $\mathbb{I}$ (σ_z) as the identity (Pauli) matrix. The third term in Hamiltonian (120) is the on-site SOI Hamiltonian. Without the presence of B the system has time-reversal symmetry and its bands are degenerate in spin throughout the whole BZ. We use the exchange field to break the time-reversal symmetry and induce a non-zero QAH effect. To identify different origins of the Chern insulator phase, the spin-orbit interaction is further decomposed into spin-conserving and spin-flip parts:

$$H_{\rm SOI} \;=\; \xi \mathbf{l}\cdot\mathbf{s} = \xi l_z s_z + \xi(l^+ s^- + l^- s^+)/2, \tag{121}$$

where $\mathbf{l}$ ($\mathbf{s}$) is the orbital (spin) angular momentum operator, and ξ is the atomic SOI strength. Since in this work we choose the direction of the spin-polarization to be aligned along z direction, the spin conserving part of the SOI, $\xi l_z s_z$, couples $\{p_x, p_y\}$ orbitals, while the spin-flip part of Eq. (73), $\xi(l^+ s^- + l^- s^+)/2$, couples p_z and $\{p_x, p_y\}$ orbitals via a flip of spin and a ± 1 change in the orbital quantum number.

We concentrate here only on the case of the p_z orbitals present around the Fermi energy, with other states being pushed away much higher in energy. This case is particularly relevant for graphene physics. For the values of the spin-orbit strength and hopping parameters we choose those of Bi from Ref. [15]. First, consider the case when only spin-conserving SOI is present. On a buckled honeycomb lattice, such as Bi bilayer or silicine, p_z orbitals can hybridize directly with the $\{p_x, p_y\}$ orbitals on the neighbring site, and complex hoppings within the sublattice can be induced via the spin-conserving part of SOC which acts between p_x and p_y states. As illustrated with a sketch in Fig. 4, in this mechanism the corresponding virtual transitions read: $|p_z^{\rm A}\uparrow\rangle \overset{t_{\rm NN}}{\rightarrow} |p_{x,y}^{\rm B}\uparrow\rangle \overset{\xi l_z s_z}{\rightarrow} |p_{x,y}^{\rm B}\uparrow\rangle \overset{t_{\rm NN}}{\rightarrow} |p_z^{\rm A}\uparrow\rangle$, where $t_{\rm NN}$ indicates the direct hopping between p_z and $p_{x,y}$ orbitals on the neighboring sites, while superscripts A and B denote the nearest neighbor atomic sites in sublattice A and B. The SOI here acts as a magnetic field which is responsible for the generation of phase ϕ within the Haldane model. If we consider only one spin, we can indeed show that effective SOI induced next-nearest-neighbor (NNN) hopping leads to the opening of the gap at the Fermi energy (of the size Δ_1 in Fig. 4(a)) and the Chern number of the spin-up bands acquires a value of $+1$. Since the SOI coefficients are complex conjugate for spin-up and spin-down electrons, the ϕ of the complex hoppings for spin-down electrons is

opposite in sign to spin-up electrons and same holds for the Chern numbers. Since the bands are spin-degenerate with $B = 0$ (we apply as small exchange field in Fig. 4 to artificially separate bands of opposite spin for visibility), this results in a zero total Chern number, and the system resides in a *topological insulator phase* (see here also manuscript A10 by Gustav Bihlmayer).

On-site spin-flip SOI can give rise to complex next nearest neighbor hopping too, even if there is no direct hybridization between p_z and $\{p_x, p_y\}$ orbitals, Fig. 4(b). In this case, the corresponding virtual transitions are: $|p_z^{\rm A}\uparrow\rangle \overset{\xi_{\rm flip}}{\rightarrow} |p_{x,y}^{\rm A}\downarrow\rangle \overset{t_{\rm NN}}{\rightarrow} |p_{x,y}^{\rm B}\downarrow\rangle \overset{t_{\rm NN}}{\rightarrow} |p_{x,y}^{\rm A}\downarrow\rangle \overset{\xi_{\rm flip}}{\rightarrow} |p_z^{\rm A}\uparrow\rangle$ where $\xi_{\rm flip} = \xi(l^+s^- + l^-s^+)/2$, and $t_{\rm NN}$ stands for the direct hybridization between $p_{x,y}$ orbitals on neighboring A and B sites. The corresponding NNN hopping is again between electrons of the same spin, owing to two spin-flip processes which take place in between. In analogy to the case with spin-conserving SOI considered previously, the effective hoppings within A and B sublattices for fixed spin are of opposite sign and the resulting gap Δ_2 which opens due to latter virtual transitions is again topologically nontrivial. The resulting non-zero Chern numbers of the degenerate without B spin-up and spin-down bands are exactly the same as previously. As in the previous case, the coupling between the spin-up and spin-down p_z bands does not occur, and without an exchange field, Fig. 4(b) corresponds exactly to the topological insulator phase in graphene [16].

How do we make Chern insulators out of systems above? In principle, if we could apply a very strong exchange field which would shift the bands of a certain spin very high up in energy, we would readily obtain a QAH insulator. In real materials this is however seldomly achievable, since the magnitudes of typical exchange fields are normally smaller than the typical band width. Thus, the only way would be in breaking the time-reversal symmetry with a finite B, and ensuring that a topologically non-trivial band gap opens at the Fermi energy where bands of opposite spin meet. For example, we could start with a situation depicted in Fig. 4(a) with a small B, and add an admixture of the spin-flip SOI to the Hamiltonian. This will open a gap at the points where the spin-up and spin-down bands were degenerate, see Fig. 4. In the vicinity of such a point the distribution of spin becomes non-trivial, as we can see in Fig. 4, namely, e.g. at the K-point the spin-distribution of the occupied band has a skyrmion structure. At this point, we can relate the distribution of spin to the distribution of vector $\mathbf{d}$ and, according to the two-band analysis presented above, we can explain the fact that the Chern number of the occupied band changes. The spin distribution of the corresponding conduction band is also a skyrmion, but with an opposite winding number, which results in the opposite change of the Chern number of this band. What we have just achieved is the exchange of the Chern number between the bands of opposite spin at the points at the Fermi energy where bands of opposite spin hybridize. We call such points *spin-mixing points*. In this particular example the spin-mixing point is the Weyl point in the space of $(k_x, k_y, \xi_{\rm flip})$, where $\xi_{\rm flip}$ is the strength of the spin-flip SOI. We have two spin-mixing Weyl points in our 3D space: (K, $\xi_{\rm flip} = 0$) and (K′, $\xi_{\rm flip} = 0$). The change in the Chern number upon going through these Weyl points can be computed based on the dispersion of the bands, and the total Chern number of all occupied states can be calculated to be -2 for a realistic situation of $\xi_{\rm flip} > 0$. We have thus achieved a QAH state.

We remark here, that Hamiltonians of real materials can be very complicated with many states present at the Fermi energy and various structural, spin-orbit and magnetic effects taking place. The phase diagrams of such materials as a function of parameters in the Hamiltonian can be studied from first principles methods. See for example the phase diagram of Bi(111) bilayer as a function of an exchange field and SOI strength calculated from *ab initio* in Fig. 3.

We would like to comment on the relation between Chern numbers and transport properties

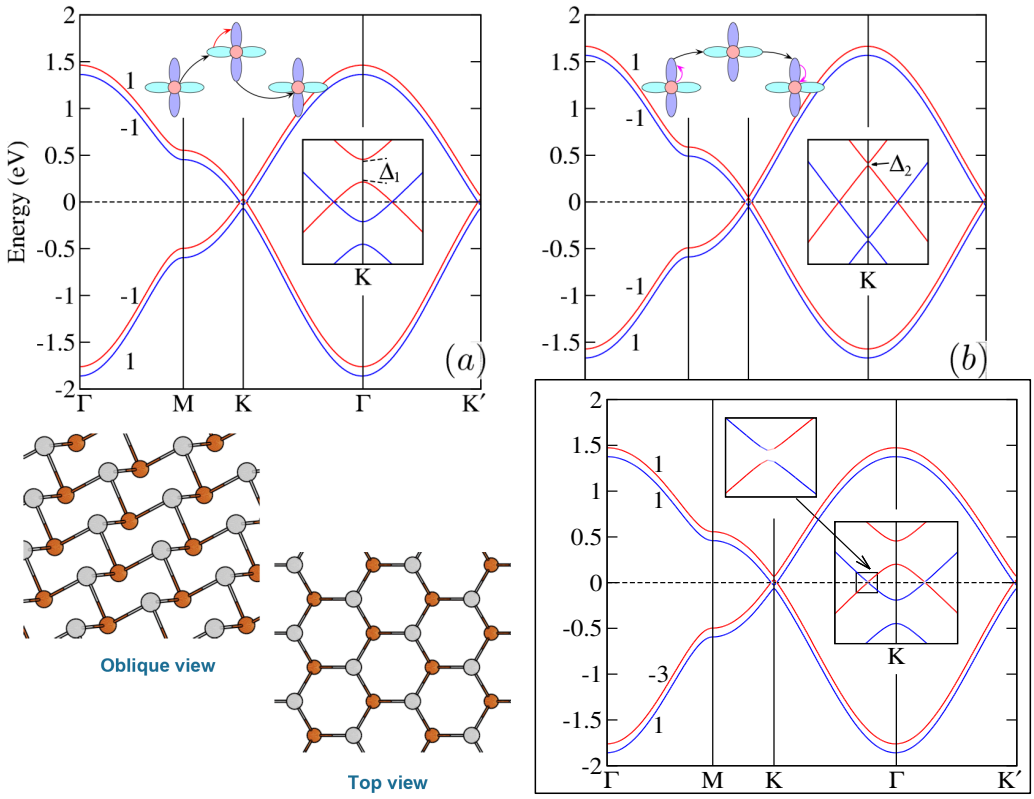


Fig. 4: *Topological analysis of p_z bands in a buckled honeycomb bilayer. A small exchange field has been applied in all cases, yielding a QAH state in lower right figure. Without an exchange field, systems in (a) and (b) would be in a topological insulator state. Left up and right up figures correspond to the case with only spin-conserving SOC, and only spin-flip SOC included, while the full SOC is considered in the lower right figure. Red (blue) stands for the spin-down (spin-up) states. Dashed horizontal lines indicate the Fermi energy. In the case of isolated bands, numbers denote the Chern number for each individual band, while for overlapping bands, numbers stand for the Chern numbers of nonhybridizing spin-up (red) and spin-down (blue) bands. Insets display the electronic structure of the Dirac point at the Fermi energy, and sketches illustrate different channels for complex nearest-neighbor hopping [red circles denote p_z orbitals, while $p_{x,y}$ orbitals are indicated by blue ellipsoids; black (red) arrows depicts the nearest-neighbor hoppings (SOC hybridization), respectively]. Lower left: top and side view of the Bi(111) bilayer. For more details see text and Ref. [15].*

of 2D insulators. We have learned by now that a Chern insulator has a quantized transverse charge conductance, proportional to the value of the Chern number. This means that e.g. for the case of Fig. 4(a) the transverse charge conductance vanishes. The carriers of each spin separately, however, possess a quantized charge conductance. In an applied electric field the carriers of opposite spin will move in opposite directions, which will lead to the generation of the transverse *spin current*. Thus obtained spin conductance is quantized since the conductance for each spin is quantized, and it is proportional to the difference between the Chern numbers for each spin. This number is called the *spin Chern number*. The system of Fig. 4(a) is a simple example of a *quantum spin Hall* (QSH) insulator, or, equivalently, a 2D topological insulator. The concept of the spin Chern number can be generalized to the cases where spin

is not conserved, and even to the case of broken time-reversal symmetry. For 2D insulators with time-inversion the spin Chern number can be used alternatively to the $\mathbb{Z}_2$ index in order to classify topological phases [5].
In metals, the Chern number can be formally calculated but it is not quantized, since for bands which cross the Fermi energy the integral of the Berry curvature goes only along the patches of the BZ where the band is occupied. Nevertheless, the so-calculated integral of the Berry curvature over all occupied states gives the value of the *intrinsic* anomalous Hall effect (AHE). In contrast to insulators, the presence of Fermi surface in metals also leads to promotion of the Hall current which comes from impurity scattering – this is the so-called *extrinsic* AHE. The relation between the magnitudes of the intrinsic and extrinsic currents strongly depends on the details of the electronic structure and disorder [13]. The theory of impurity scattering will be considered in detail by Phivos Mavropoulos in manuscript A5 of this book. In analogy to the relation between the Chern and spin Chern numbers, if the spin is conserved, the anomalous Hall conductivity can be decomposed into a sum of contributions coming from spin-up and spin-down bands. The difference between the two is proportional to the value of the spin Hall conductivity, that is, it is related to the magnitude of the transverse spin current caused by an electric field. This effect is called the *spin Hall effect* (SHE).

3.4 Dynamics of wavepackets in solids

A very powerful approach to study the properties of solids lies in rewriting the problem in terms of so-called wavepackets. The general description of the electron dynamics in external fields and perturbations can be rigorously provided referring to semiclassical dynamics of such wavepackets. The wavepackets are obtained by a convolution of Bloch states with the envelope function which is centered around a certain $\mathbf{k}$-vector. The wavepackets are made such that they are localized both in reciprocal and real space, and the evolution of their center of mass in both spaces can be directly related to the transport chracteristics of a solid. The quantum mechanical description of wavepackets is an intricate science on its own, and here, we will only provide intuitive arguments which can be used to understand how the exact equations of wavepacket motion are obtained [6].

Uniform electric field.
For simplicity, let us first consider the effect of an electric field $\mathbf{E}$ present in a solid. Let us say that without the field the unperturbed Hamiltonian of the system looks like

$$H(t=0) = \frac{\mathbf{p}^2}{2} + V(\mathbf{r}), \tag{122}$$

with corresponding Bloch vectors $\mathbf{q}$ and the crystal momentum representation $H_{\mathbf{q}}$ of the Hamiltonian. To model the effect of the uniform electric field, we apply a constant in space, but varying in time vector potential $\mathbf{A}(t)$ such that $-\frac{\partial \mathbf{A}(t)}{\partial t} = \mathbf{E}$. This modifies the corresponding lattice Hamiltonian as follows:

$$H(t>0) = \frac{1}{2}\left(\mathbf{p} + \mathbf{A}(t)\right)^2 + V(\mathbf{r}). \tag{123}$$

Since the constant in space vector potential does not break the periodicity of the crystal, it cannot couple the unperturbed wavefunctions with different values of $\mathbf{q}$ and it changes the energy of the states with an overall constant, which can be ignored as we shall see later. Therefore, once we are looking at a state labeled with a certain value of $\mathbf{q}$, during the evolution of this state $\dot{\mathbf{q}} = 0$.

Nevertheless, we can also number our state in terms of the $\mathbf{k}$ vector, which is the "proper" Bloch vector of the Hamiltonian $H_{\mathbf{k}}(t > 0)$, and which is called *gauge-invariant momentum*. The relation between the $\mathbf{k}$ and $\mathbf{q}$ reads as follows:

$$\mathbf{k} = \mathbf{q} + \mathbf{A}(t) = \mathbf{k}(t), \tag{124}$$

that is, the wavefunction at $\mathbf{k}(t)$ which solves the Schrödinger equation for $H(t)$, is identical to the wavefunction which solves the Schrödinger equation for $H(t = 0)$ but at a wavevector $\mathbf{q} = \mathbf{k} - \mathbf{A}(t)$. From the latter equation, it follows that

$$\dot{\mathbf{k}} = -\mathbf{E}. \tag{125}$$

In order to employ the expression for the velocity of a certain state, we write the time-dependence of the Hamiltonian as follows: $H^t_{\mathbf{k}} = H_{\mathbf{k}(t)}$, which leads to $\partial_t H^t_{\mathbf{k}} = \partial_{\mathbf{k}} H_{\mathbf{k}} \cdot \partial_t \mathbf{k}(t) = -\mathbf{E} \cdot \partial_{\mathbf{k}} H_{\mathbf{k}}$. Substituting the latter into the expression for the velocity (107) with t playing the role of λ, we get:

$$\dot{\mathbf{r}} := v_{n\mathbf{k}} = \partial_{\mathbf{k}} \varepsilon_{n\mathbf{k}} - \mathbf{E} \times \mathbf{\Omega}_n(\mathbf{k}). \tag{126}$$

From these equations we can easily understand the Berry phase origin of the intrinsic anomalous Hall effect in ferromagnets, as well as the expression for the Hall conductance in terms of the Berry curvature given by Eq. (111), taking into account that the contribution from the group velocity when integrated over the BZ, vanishes.

Uniform magnetic field.
Let us assume now a situation in which we have applied an external uniform magnetic field $\mathbf{B} = \mathrm{curl}\mathbf{A}(\mathbf{r})$. This situation is analogous to the previously considered case only with $\mathbf{A}(\mathbf{r})$ replacing the $\mathbf{A}(t)$, but the way of treating the two situations is quite different. Namely, generally speaking, the vector potential $\mathbf{A}$ breaks the periodicity of the lattice. This obstacle can be overcome by assuming that the vector potential varies very slowly in space, so that locally at a given point $\mathbf{R}$ in space the periodicity of the lattice is preserved and the "local" Bloch momentum $\mathbf{k}$ is well-defined. Seen from point $\mathbf{R}$, the presence of the vector potential is then just a constant shift of the momentum operator, analogously to the case of the electric field. The time-dependent process associated with this assumption is the propagation of an electron wavepacket, with the center in real space at $\mathbf{R}(t)$, through a slowly varying medium with the Hamiltonian $H = H(\mathbf{k}(\mathbf{R}(t))$, where the gauge-invariant momentum $\mathbf{k}$ is given by:

$$\mathbf{k} = \mathbf{q} + \mathbf{A}(\mathbf{R}) = \mathbf{k}(t) = \mathbf{k}(\mathbf{R}(t)). \tag{127}$$

Following the same logic as previously, that is that $\dot{\mathbf{q}} = 0$, we derive the equation of propagation in the reciprocal space:

$$\dot{\mathbf{k}} = \mathbf{B} \times \dot{\mathbf{R}} = \mathbf{\Omega}_{\mathbf{R}} \times \dot{\mathbf{R}}, \tag{128}$$

where $\mathbf{\Omega}_{\mathbf{R}}$ is real space Berry curvature (37), which corresponds to the magnetic field, as we saw from the consideration of the AB-effect. On the other hand, since $\partial_t H^t_{\mathbf{k}} = \partial_{\mathbf{k}} H_{\mathbf{k}} \cdot \dot{\mathbf{k}}$, the equation of motion of the wavepacket generalizes to

$$\dot{\mathbf{R}} = \partial_{\mathbf{k}} \varepsilon_{n\mathbf{k}} + \mathbf{\Omega}_n(\mathbf{k}) \times \dot{\mathbf{k}}. \tag{129}$$

Explicit dependence on $\mathbf{R}$.
Imagine now that the dependence of the local Hamiltonian on the slowly varying spatial coordinate $\mathbf{R}$ is not only via the vector potential, but also via the crystal potential $V(\mathbf{R})$, or

($\mathbf{R}$-dependent) exchange coupling of the spin of a propagating electron to the ($\mathbf{R}$-dependent) magnetic texture (the case we will consider in detail in the next section). Namely, let us suppose that the Hamiltonian assumes a more general dependence $H = H(\mathbf{k}(\mathbf{R}(t)), \mathbf{R}(t))$. In this case the time-derivative of the Hamiltonian

$$\partial_t H(\mathbf{k}(\mathbf{R}(t)), \mathbf{R}(t)) = \partial_{\mathbf{k}} H_{\mathbf{kR}} \cdot \dot{\mathbf{k}} + \partial_{\mathbf{R}} H_{\mathbf{kR}} \cdot \dot{\mathbf{R}}. \tag{130}$$

This leads to the fact that the velocity of a wavepacket through the $\mathbf{R}$-texture acquires an additional contribution due to the *mixed Berry curvature* $\Omega_{\mathbf{kR}}$, which can be expressed in terms of the derivatives of $u_{\mathbf{kR}}$ with respect to $\mathbf{k}$ and with respect to $\mathbf{R}$. This can be intuitively understood by looking at Eq. (107), into which both of these derivatives will enter upon the time evolution of the states. We previously defined the mixed Berry curvature in the context of electric polarization, for which $\mathbf{R}$ was replaced with λ.

General case.

While the hand-waving arguments we provided to derive the expression for the velocity of a wavepacket propagating through an $\mathbf{R}$-texture are qualitatively correct, the rigorous quantum mechanical derivation of the equations of motion of the center of a wavepacket in the $(\mathbf{R}, \mathbf{k}, t)$ phase space has been given by Sundaram and Niu in Ref. [17]. If we have an explicit time-dependence of the Hamiltonian (e.g. via a moving magnetic texture), we should assume the general dependence of the Hamiltonian $H = H(\mathbf{k}, \mathbf{R}, t)$, given that the texture varies very slowly in space and in time. In the latter expression for the Hamiltonian, $\mathbf{k}$ stands for the "local" Bloch vector, which is a good quantum number under the assumption that locally around the point $\mathbf{R}$ in space the texture can be approximated as a constant and the lattice periodicity is preserved. In terms of preceeding subsections $\mathbf{k}$ plays the role of $\mathbf{q}$. Correspondingly, all wavefunctions acquire additional dependence on $\mathbf{R}$ and t: $u_{n\mathbf{k}} \longrightarrow u^t_{n\mathbf{kR}}$. In accordance to this additional dependence, we can introduce the following gauge potentials:

$$\mathcal{A}^n_t = i \left\langle u^t_{n\mathbf{kR}} \,\middle|\, \partial_t u^t_{n\mathbf{kR}} \right\rangle, \quad \mathcal{A}^n_{\mathbf{k}} = i \left\langle u^t_{n\mathbf{kR}} \,\middle|\, \partial_{\mathbf{k}} u^t_{n\mathbf{kR}} \right\rangle, \quad \mathcal{A}^n_{\mathbf{R}} = i \left\langle u^t_{n\mathbf{kR}} \,\middle|\, \partial_{\mathbf{R}} u^t_{n\mathbf{kR}} \right\rangle, \tag{131}$$

which can be all seen as one gauge potential $\mathcal{A}^n$ on the $(\mathbf{R}, \mathbf{k}, t)$-manifold with components given above. The corresponding curvature of our phase space $\Omega^n_{ij} = \partial_i \mathcal{A}^n_j - \partial_j \mathcal{A}^n_i$ has then corresponding components:

$$\Omega^n_{\mathbf{k}_i\mathbf{k}_j} = -2\mathrm{Im} \left\langle \partial_{\mathbf{k}_i} u^t_{n\mathbf{kR}} \,\middle|\, \partial_{\mathbf{k}_j} u^t_{n\mathbf{kR}} \right\rangle, \quad \Omega^n_{\mathbf{R}_i\mathbf{R}_j} = -2\mathrm{Im} \left\langle \partial_{\mathbf{R}_i} u^t_{n\mathbf{kR}} \,\middle|\, \partial_{\mathbf{R}_j} u^t_{n\mathbf{kR}} \right\rangle, \tag{132}$$

and the mixed curvature:

$$\Omega^n_{\mathbf{k}_i\mathbf{R}_j} = -2\mathrm{Im} \left\langle \partial_{\mathbf{k}_i} u^t_{n\mathbf{kR}} \,\middle|\, \partial_{\mathbf{R}_j} u^t_{n\mathbf{kR}} \right\rangle. \tag{133}$$

The expressions for the time-involving components of the curvature can be written analogously. The equations of motion of the center of the wavepacket which is centered at points $\mathbf{R}$ and $\mathbf{k}$ in real and reciprocal space, respectively, then read:

$$\dot{\mathbf{R}} = \partial_{\mathbf{k}} \mathcal{E}^t_{n\mathbf{Rk}} - \left(\Omega^n_{\mathbf{kR}} \cdot \dot{\mathbf{R}} + \Omega^n_{\mathbf{kk}} \cdot \dot{\mathbf{k}} \right) - \Omega^n_{\mathbf{k}t}, \tag{134}$$

$$\dot{\mathbf{k}} = -\partial_{\mathbf{R}} \mathcal{E}^t_{n\mathbf{Rk}} + \left(\Omega^n_{\mathbf{RR}} \cdot \dot{\mathbf{R}} + \Omega^n_{\mathbf{Rk}} \cdot \dot{\mathbf{k}} \right) - \Omega^n_{\mathbf{R}t}. \tag{135}$$

Note that the band energy $\varepsilon^t_{n\mathbf{kR}}$ acquires an additional contribution:

$$\mathcal{E}^t_{n\mathbf{Rk}} = \varepsilon^t_{n\mathbf{kR}} + \delta\varepsilon^t_{n\mathbf{kR}} = \varepsilon^t_{n\mathbf{kR}} - \mathrm{Im} \sum_i \left\langle \partial_{\mathbf{R}_i} u^t_{n\mathbf{kR}} \,\middle|\, \varepsilon^t_{n\mathbf{kR}} - H(\mathbf{k}, \mathbf{R}, t) \,\middle|\, \partial_{\mathbf{k}_i} u^t_{n\mathbf{kR}} \right\rangle. \tag{136}$$

In the context of situations considered previously, it is clear that the $\Omega^n_{\mathbf{RR}}$ curvature plays the role of the real-space magnetic field, while the $\Omega^n_{\mathbf{kk}}$ is the curvature in reciprocal space which gives rise to the anomalous Hall effect. When the system is subject to constant electric and magnetic fields, the mixed Berry curvature is zero, and equations of motion for the wavepackets are reduced to:

$$\dot{\mathbf{R}} = \partial_{\mathbf{k}}\mathcal{E}_{n\mathbf{k}} - \dot{\mathbf{k}} \times \mathbf{\Omega}^n_{\mathbf{kk}} \tag{137}$$

$$\dot{\mathbf{k}} = -\mathbf{E} - \dot{\mathbf{R}} \times \mathbf{B}, \tag{138}$$

where $\mathbf{k}$ stands now for gauge-invariant momentum from before: $\mathbf{k} = \mathbf{q} + \mathbf{A}(\mathbf{R})$. If we use now that $\partial_{\mathbf{R}} = (\partial_{\mathbf{R}}\mathcal{A})\partial_{\mathbf{k}}$, we can write down the contribution to the energy in terms of $\mathbf{k}$-derivatives only:

$$\mathcal{E}_{n\mathbf{k}} = \varepsilon_{n\mathbf{k}} - \mathbf{B} \cdot \mathrm{Im}\,\langle \partial_{\mathbf{k}} u_{n\mathbf{k}} \,|\, \times [\varepsilon_{n\mathbf{k}} - H(\mathbf{k}, t)] \,|\, \partial_{\mathbf{k}} u_{n\mathbf{k}} \rangle = \varepsilon_{n\mathbf{k}} - \mathbf{B} \cdot \mathbf{m}(\mathbf{k}). \tag{139}$$

In the last equation, $\mathbf{m}(\mathbf{k})$ is given by:

$$\mathbf{m}(\mathbf{k}) = \mathrm{Im}\,\langle \partial_{\mathbf{k}} u_{n\mathbf{k}} \,|\, \times [\varepsilon_{n\mathbf{k}} - H(\mathbf{k}, t)] \,|\, \partial_{\mathbf{k}} u_{n\mathbf{k}} \rangle\,, \tag{140}$$

and it stands for the orbital moment of the wavepacket corresponding to the "internal" degree of freedom with respect to the rotation around its own axis. The last term in Eq. (139) thus corresponds to the interaction of the magnetic field with the orbital moment of the wavepacket due to rotation around its own axis.

Geometrical meaning.
Let us for simplicity drop the explicit time dependence in Eqs. (134-135). Then the equations which govern the dynamics in the phase space $\mathbf{x} = (\mathbf{k}, \mathbf{R})$ can be written in the following form:

$$(\Omega^n - J)\dot{\mathbf{x}} = \partial_{\mathbf{x}}\mathcal{E}_n \quad \Longleftrightarrow \quad \omega^n_{\alpha\beta}\dot{\mathbf{x}}_\beta = \partial_{\mathbf{x}_\alpha}\mathcal{E}_n \tag{141}$$

where for simplicity we dropped $\mathbf{R}$ and $\mathbf{k}$ indices, $\omega^n = (\Omega^n - J)$ and $J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$. These equations are nothing else but the equations of the classical Hamiltonian dynamics written in terms of *non-canonical* variables $\mathbf{x}$ for the Hamiltonian of the form $h(\mathbf{x}) = \mathcal{E}_n(\mathbf{x})$. If we introduce the Poisson bracket between two functions $f(\mathbf{x})$ and $g(\mathbf{x})$ as follows:

$$\{f, g\} := \partial_{\mathbf{x}} f^T \cdot (\omega^n)^{-1} \cdot \partial_{\mathbf{x}} g, \tag{142}$$

the Hamilton equations acquire the standard form:

$$\dot{\mathbf{x}} = \{\mathbf{x}, h\}. \tag{143}$$

The fact that the position and momentum of the wavepacket are non-canonical variables comes with a price. Namely, the canonical Hamilton equations, obtained with $\omega = -J$, satisfy the property called the *Liouville's theorem*, which states that during the Hamiltonian dynamics the volume of a certain region in phase space, say, $d\mathbf{k}d\mathbf{R}$, is preserved. It is very easy to check, that, taking initially an infinitesimal volume $d\mathbf{k}d\mathbf{R}$ and evoluting it in accordance to (141) will violate the Liouville's theorem. The reason for this lies in the fact that $\mathbf{k}$ and $\mathbf{R}$ are non-canonical variables, when the curvature form is non-zero. Mathematically speaking, we endow

our phase space with a non-canonical simplectic two-form $\omega^n = (1/2)\omega^n_{\alpha\beta} d\mathbf{k}_\alpha \wedge d\mathbf{R}_\beta$, which influences the measure of our $\mathbf{x}$-space. It is straightforward to see that the infinitesimal volume $d\mathbf{k}d\mathbf{R}$ multiplied with the function, called *modified density of states*,

$$D^n(\mathbf{x}) = \frac{1}{(2\pi)^d}\sqrt{\det \omega^n} \tag{144}$$

satisfies the Liouville's theorem. In case of canonical variables the modified density of states (DOS), $D(\mathbf{x}) = 1/(2\pi)^d$ (multiplied with the Fermi distribution function). Let us take a detailed look at the explicit expression for the case of the (generally $\mathbf{R}$-dependent) magnetic field in the system. The Poisson brackets between the $\mathbf{R}$ and $\mathbf{k}$ coordinates in this case read:

$$\{\mathbf{R}_i, \mathbf{R}_j\} = \frac{\varepsilon_{ijl}\mathbf{\Omega}^n_{\mathbf{kk},l}}{1 + \mathbf{B}\cdot\mathbf{\Omega}^n_{\mathbf{kk}}}, \quad \{\mathbf{k}_i, \mathbf{k}_j\} = -\frac{\varepsilon_{ijl}\mathbf{B}_l}{1 + \mathbf{B}\cdot\mathbf{\Omega}^n_{\mathbf{kk}}}, \quad \{\mathbf{R}_i, \mathbf{k}_j\} = \frac{\delta_{ij} + \mathbf{B}_i\mathbf{\Omega}^n_{\mathbf{kk},j}}{1 + \mathbf{B}\cdot\mathbf{\Omega}^n_{\mathbf{kk}}}, \tag{145}$$

while the determinant of ω gives:

$$\sqrt{\det \omega^n} = 1 + \mathbf{B}\cdot\mathbf{\Omega}_{\mathbf{kk}} \quad \Longrightarrow \quad D^n(\mathbf{x}) = \frac{1}{(2\pi)^d}(1 + \mathbf{B}\cdot\mathbf{\Omega}_{\mathbf{kk}}) \tag{146}$$

The volume element $D^n(\mathbf{x})d\mathbf{k}d\mathbf{R}$ has constant density of quantum states and it has to be used for computing the expectation values of the observables obtained by an integration over the phase space. We have to understand that, of course, coordinates $\mathbf{k}$ and $\mathbf{R}$ can be made canonical, although possibly less appealing physically. For such canonical variables, say, $\tilde{\mathbf{k}}$ and $\tilde{\mathbf{R}}$, the modified DOS reads $D^n(\tilde{\mathbf{x}}) = 1$ and the very same expectation values can be obtained by a direct evaluation of the integral with the volume element measure $d\tilde{\mathbf{k}}d\tilde{\mathbf{R}}$. The situation here is analogous to switching between integrals of a function in $\mathbb{R}^3$ performed for example in cartesian and spherical coordinates. The modified DOS for the non-canonical variables is thus analogous to the Jacobian of the transformation between cartesian and spherical coordinates. Below we briefly outline the consequences of the non-canonicity of the $\mathbf{R}$ and $\mathbf{k}$ coordinates for a selected number of physical properties of the system.

Fermi volume.
The change in the density of electronic states in the $(\mathbf{R}, \mathbf{k})$ space inevitably changes the expectation values of quantum operators which are obtained as integrals over $(\mathbf{R}, \mathbf{k})$ space. Consider the Fermi volume which a given number of electrons occupies. The number of occupied electrons, N_e, is given by:

$$N_e = \int_{-\infty}^{E_F} \frac{d\mathbf{k}}{(2\pi)^d}(1 + \mathbf{B}\cdot\mathbf{\Omega}_{\mathbf{kk}}). \tag{147}$$

In case of a small magnetic field, keeping the number of electrons in an insulator constant, we come to the change in the Fermi volume by:

$$\Delta V_F = -\int_{\mathrm{BZ}} \frac{d\mathbf{k}}{(2\pi)^d}\,\mathbf{B}\cdot\mathbf{\Omega}_{\mathbf{kk}}. \tag{148}$$

In case of two dimensions, we therefore get that $\Delta V_F = \sigma_{xy}\cdot B$.

Quantum Hall conductance.
Thermodynamically, the change in the free energy of the system can be written as:

$$dF = -\mathbf{M}_L\, d\mathbf{B} - N_e\, d\mu - S\, dT, \tag{149}$$

where $\mathbf{M}_L$ is the magnetization induced by the magnetic field $\mathbf{B}$, μ is the chemical potential, S is entropy and T is temperature. Streda formula states that the Hall conductivity of a two-dimensional finite sample is the derivative of the electron number at a certain chemical potential and temperature with respect to an applied field:

$$\sigma_{xy} = -\left(\partial N_e/\partial \mathbf{B}\right)_{\mu,T}. \tag{150}$$

Using the expression derived previously for N_e, we obtain the known expresion for the σ_{xy} in therms of the Berry curvature in $\mathbf{k}$-space. The Streda formula can be understood intuitively by thinking in terms of a time dependent rise of the magnetic field in some region, which will generate an electric field along the boundary of this region. This in turn will cause the "leakage" of the charge from the region due to the anomalous Hall effect.

Orbital magnetization.
The correction to the density of states allows us also to derive an explicit expression for the orbital magnetization from (149), since it is defined as $\mathbf{M}_L = -(\partial F/\partial \mathbf{B})_{\mu,T}$. The expression for $\mathbf{M}_L$ can be derived referring to the explicit expression for F ($\beta = 1/k_B T$, and we drop the summation over bands for simplicity):

$$F = -\frac{1}{\beta(2\pi)^d}\int d\mathbf{k}\,(1 + \mathbf{B}\cdot\mathbf{\Omega}_{\mathbf{kk}})\ln\left(1 + e^{-\beta(\mathcal{E}_{n\mathbf{k}}-\mu)}\right). \tag{151}$$

By differentiating this expression with respect to the magnetic field, we obtain:

$$\mathbf{M}_L = \int_{\mathrm{BZ}}\frac{d\mathbf{k}}{(2\pi)^d}\,f(\mathbf{k})\mathbf{m}(\mathbf{k}) + \int_{\mathrm{BZ}}\frac{d\mathbf{k}}{(2\pi)^d}\,\mathbf{\Omega}_{\mathbf{kk}}\ln\left(1 + e^{-\beta(\varepsilon_{n\mathbf{k}}-\mu)}\right), \tag{152}$$

where f is the Fermi occupation function. From this expression it is clear that two effects contribute to the total orbital magnetization in a solid. The first one comes from the $\mathbf{m}(\mathbf{k})$ correction to the energy, which can be identified with the orbital moment of a wavepacket as it rotates around its axis. The second contribution comes from the center-of-mass motion of the wavepackets, and can be expressed in terms of the anomalous Hall conductivity. This latter contribution is thus a direct consequence of the modification in the density of states in the phase space due to its non-trivial structure as expressed in terms of the curvature. At zero temperature, the expression for the orbital magnetization thus reads:

$$\mathbf{M}_L = \sum_n \int_{\mathrm{BZ}}\frac{d\mathbf{k}}{(2\pi)^d}\,f_{n\mathbf{k}}\left[\mathbf{m}_n(\mathbf{k}) + (\varepsilon_{n\mathbf{k}} - \mu)\mathbf{\Omega}^n_{\mathbf{kk}}\right]. \tag{153}$$

The expression for the quantized Hall conductivity in an insulator can be rederived using the Maxwell relation $(\partial \mathbf{M}_L/\partial\mu)_{\mathbf{B},T} = (\partial N_e/\partial \mathbf{B})_{\mu,T}$. The latter expression suggests that in a Chern insulator the orbital magnetization varies linearly with the chemical potential. The mechanism for this effect lies in the presence of metallic states at the boundary of a Chern insulator.

3.5 Topological Hall effect

Emergent field and topological Hall effect.
The term topological Hall effect (THE) is normally referred to in the context of a magnetic

material which exhibits a spatial variation of the magnetization $\mathbf{M}(\mathbf{r})$. The simplest possible Hamiltonian in this case, which gives rise to non-trivial effects, reads:

$$H = \frac{\mathbf{p}^2}{2} - J\boldsymbol{\sigma} \cdot \mathbf{M}(\mathbf{r}), \tag{154}$$

where J is the coupling constant, $\boldsymbol{\sigma}$ is the vector of Pauli matrices, and we shall assume that the magnitude of the magnetization is constant, that is, $\mathbf{M}(\mathbf{r}) = M\mathbf{n}(\mathbf{r}), |\mathbf{n}(\mathbf{r})| = 1$. How do we attack the problem of finding the transport properties of such a system? The brute force method suggests that we take a finite sample, diagonalize the Hamiltonian, find the spectrum and wavefunctions, and calculate the required properties. If the magnetization texture exhibits a periodic lattice structure, we can even apply the Bloch apparatus to arrive at the spectrum $E_{n\boldsymbol{\varkappa}}$ and wavefunctions $\Psi_{n\boldsymbol{\varkappa}}$, where $\boldsymbol{\varkappa}$ is the Bloch vector from the reciprocal space. How does the Berry phase enter into this picture?

Practically, the Berry phase viewpoint at the problem is motivated by the consideration that very often the typical scale of the variation of the texture in real space is very large, which makes the brute force approach cumbersome. In this way we assume that locally around a point $\mathbf{R}$ in real space the magnetization direction is constant, resulting in a set of eigenvalues and wavefunctions of the "local" Hamiltonian, $\varepsilon_{n\mathbf{R}}$ and $\psi_{n\mathbf{R}}$, respectively. The properties of the system are then related to the dynamics of electrons as they move through the texture, or, in other terms, as parameter $\mathbf{R}$ is varied $-$ the Hamiltonian is then seen as a parametrized Hamiltonian $H(\mathbf{R})$. This is the typical setup in which Berry phase physics arises.

Let us analyze the situation of Eq. (154) from the standpoint of sections 2.4 and 3.3. We introduce a mapping (113), $\chi : \mathbb{R}^3 \to S^2$, which maps the direction of $\mathbf{n}$ at point $\mathbf{R} \in \mathbb{R}^3$ to a point on S^2, which can be ascribed angles θ and φ. This mapping is analogous to that given by (113), where the $\mathbf{k}$-space is replaced with $\mathbf{R}$-space. We are now to study the *adiabatic* dynamics of a wavefunction which solves Eq. (154) as it follows a certain trajectory $\mathcal{C}$ in real space, or, equivalently, a corresponding trajectory $\chi(\mathcal{C})$ on S^2, see Fig. 1. In spin space, at each point on S^2, as in the case of sections 2.4 and 3.3, we have two eigenenergies $\varepsilon_{\uparrow(\downarrow)} = \mp JM$ (assuming for simplicity that kinetic energy is zero), and wavefunctions $\psi_{\uparrow(\downarrow)}(\mathbf{R})$ which solve (154) for corresponding $\mathbf{R}$. The condition of the adiabaticity, as we defined it previously, requires that the time scale of the evolution of the wavefunction along $\mathcal{C}$ is much smaller than the typical time-scale of the transitions which cause spin-flip events. This corresponds to the situation of remaining within the same spin-up or spin-down subband during the dynamics. In this case the situation considered here can be made conceptually and technically analogous to that we considered for Chern insulators in section 3.3 by replacing $\mathbf{R}$ with $\mathbf{k}$.

As we recall from section 2.4, the problem posed on S^2 accounts to two independent copies of the Dirac monopole with the charge $+1$ and -1 for spin-up and spin-down electrons, respectively. Each spin σ endows S^2 with the Berry connection $\mathcal{A}^\sigma_{\mathbf{n}}$ defined on northern and southern hemispheres. The corresponding curvature $\Omega^\sigma_{\mathbf{n}}$ has components which correspond to the field of a magnetic monopole at the origin, according to (49), and has opposite sign for up and down electrons, so that summed up over both states it gives zero. Then the Berry phase $\gamma(\mathcal{C})$ an electron of certain σ picks up when it travels along $\mathcal{C}$ can be computed from the corresponding Berry phase of the path $\chi(\mathcal{C})$. The connection between the two Berry phases can be computed by writing down the relation between the Berry curvatures in both spaces. Namely, $\Omega^\sigma_{\mathbf{n}}$ and

$$\Omega^\sigma_{\mathbf{R}} \equiv \Omega^{ij,\sigma}_{\mathbf{R}\mathbf{R}} = -2\mathrm{Im}\,\langle \partial_{\mathbf{R}_\mathrm{i}}\psi_\sigma(\mathbf{R}) | \partial_{\mathbf{R}_\mathrm{j}}\psi_\sigma(\mathbf{R})\rangle \tag{155}$$

are connected in a way similar to that we considered for Chern insulators, Eq. (114):

$$\Omega^{ij,\sigma}_{\mathbf{R}} = \sigma\, \mathbf{n} \cdot \left(\partial_{\mathbf{R}_i}\mathbf{n} \times \partial_{\mathbf{R}_j}\mathbf{n}\right)/2 = \mathbf{\Omega}^{\sigma}_{\mathbf{n}} \cdot \left(\partial_{\mathbf{R}_i}\mathbf{M} \times \partial_{\mathbf{R}_j}\mathbf{M}\right). \tag{156}$$

According to Eqs. (134)-(135) the $\Omega^{\sigma}_{\mathbf{R}}$ curvature will insert a Lorentz force on a moving through the texture electron leading to transverse current, in complete analogy to the ordinary Hall effect in presence of an external magnetic field. It thus makes sense to think of $\mathbf{\Omega}^{\sigma}_{\mathbf{R}}$ as an *emergent* magnetic field $\mathbf{B}^{\sigma}_{e}$. In contrast to the ordinary Hall effect, the Hall effect due to $\mathbf{B}^{\sigma}_{e}$ is opposite for opposite spins, and it is called the *topological Hall effect*. The topological Hall effect is a purely geometrical phenomenon in that it is driven by (adiabatically slow) modulation of the magnetic texture in space, and in that it can be written in terms of Berry curvature of spin-up and spin-down electrons, which sums up to zero when both spins are occupied.
The magnitude of the emergent magnetic field at given point $\mathbf{R}$ directly depends on how quickly the texture changes in real space, according to Eq. (156). In case of a texture which exhibits a periodic modulation in space, such as e.g. in the case of a skyrmion lattice in MnSi, it is useful to define the averaged over the magnetic unit cell emergent magnetic field. From the analogy between the real-space Berry curvature and curvature of the Dirac monopole in $\mathbf{n}$-space, we can see that the integral of the emergent field over the unit cell is proportional to the solid angle which the vector $\mathbf{n}$ "draws" on a sphere as $\mathbf{R}$ is varied in the unit cell. If the texture is such that $\mathbf{n}$ covers the sphere completely, then the integral emergent field is proportional to 4π, and the constant of proportionality is the integer number of times that $\mathbf{n}$ winds around the sphere. This is the so-called *winding number* we also mentioned previosly for Chern insulators. For example, in case of a skyrmionic lattice in MnSi, the winding number is -1. Since in MnSi the skryrmion lattice constant is approximately 165 Å, the magnitude of the emergent magnetic field $B^{\downarrow}_{e}$ can be estimated to be -13 T [18].

Topological Hall effect in terms of a ferromagnetic medium.
Let us now try to rewrite everything from the point of view of an electron which follows the direction of the magnetization of the texture but in the reference frame of the magnetization. In adiabatic approximation, the expectation value of $\boldsymbol{\sigma}$ evaluated on the wavefunction $\psi_\sigma(\mathbf{R})$ is aligned along $\mathbf{M}(\mathbf{R})$. We can thus apply the transformation $U^\dagger(\mathbf{R})$, inverse to the one given by Eq. (42): $\psi_\sigma(\mathbf{R}) \longrightarrow U^\dagger(\mathbf{R})\psi_\sigma(\mathbf{R}) = \psi'_\sigma(\mathbf{R})$, which is proportional to $|\sigma \mathbf{e}_3\rangle$ from (43), i.e., the spin part of $\psi'_\sigma(\mathbf{R})$ is either parallel or antiparallel to z-axis depending on the spin. We want to ask a question: if $\psi_\sigma(\mathbf{R})$ solves the Shrödinger equation (3) with $H(\mathbf{R})$ given by (154), what is the Hamiltonian $H'(\mathbf{R})$ for which $\psi'_\sigma(\mathbf{R})$ is the solution of the corresponding Schrödinger equation? The answer can be simply obtained by substituting the $U(\mathbf{R})\psi'_\sigma(\mathbf{R}) = \psi_\sigma(\mathbf{R})$ into the Schrödinger equation for $\psi_\sigma(\mathbf{R})$, yielding:

$$i\frac{\partial \psi'_\sigma(t)}{\partial t} = \left[U^\dagger(t)H(t)U(t) - iU^\dagger(t)\frac{\partial U(t)}{\partial t}\right]\psi'_\sigma(t) = H'(t)\psi'_\sigma(t). \tag{157}$$

The Hamiltonian H' consists of two parts. The first part is neither t nor $\mathbf{R}$ dependent, since it is obviously given by:

$$H_0 = \frac{\mathbf{p}^2}{2} - JM\sigma_z, \tag{158}$$

and it provides thus a uniform ferromagnetic background on top of which the evolution of ψ'_σ takes place. The second part of the Hamiltonian H' is the one which leads to the non-trivial

Berry phase. How does it happen? Within the adiabatic approximation, the wavefunction ψ'_σ can be represented as:

$$\psi'_\sigma(t) = \psi'_\sigma(\mathbf{R}(t)) = c_\sigma(t)\,|\sigma \mathbf{e}_3\rangle\,. \tag{159}$$

The $c_\sigma(t)$ coefficient thus gives us the phase of the time-evolution, which consists of the part due to local collinear Hamiltonian H_0 (159), and of the second part due to the $-iU^\dagger \frac{\partial U}{\partial t}$ term of H'. This can be seen by substituting the ansatz for ψ' into the corresponding Schrödinger equation. Generally speaking, U and $U^\dagger$ are matrices in spin space, which will lead to coupling between the two spin channels. Within the adiabatic approximation we are going to neglect this coupling, which allows us to consider the scalar quantities $-iU^\dagger_\sigma \frac{\partial U_\sigma}{\partial t}$, given by the diagonal components of $-iU^\dagger \frac{\partial U}{\partial t}$, and opposite in sign for opposite σ. The contribution of the magnetization chirality to $c_\sigma(t)$, which we denote by $c^U_\sigma(t)$, is given by:

$$c^U_\sigma(t) = \exp\left(i\int_0^t iU^\dagger_\sigma(\tau)\frac{\partial U_\sigma(\tau)}{\partial \tau}d\tau\right). \tag{160}$$

From the latter expression it follows that if we start our evolution from some point $\mathbf{R}_0$ at $t = 0$ and move our function to $\mathbf{R}(t)$, which is connected with $\mathbf{R}_0$ by path $\mathcal{C}$, the spin part of the evoluted wavefunction, with the direction of spin being at all time either parallel or antiparallel to the z-axis, can be written as:

$$\psi'_\sigma(\mathbf{R}) = \exp\left(i\int_0^t iU^\dagger_\sigma(\tau)\frac{\partial U_\sigma(\tau)}{\partial \tau}d\tau\right)\psi'_\sigma(\mathbf{R}_0) = \exp\left(i\int_{\mathbf{R}_0}^{\mathbf{R}} \boldsymbol{\mathcal{A}}^\sigma(\mathbf{R})\,d\mathbf{R}\right)\psi'_\sigma(\mathbf{R}_0), \tag{161}$$

which is equivalent to a situation of the AB-effect from section 2.3, with the vector potential given by:

$$\boldsymbol{\mathcal{A}}^\sigma(\mathbf{R}) = iU^\dagger_\sigma(\mathbf{R})\nabla_{\mathbf{R}}U_\sigma(\mathbf{R}). \tag{162}$$

Written explicitly by referring to (42) the vector potential reads:

$$\boldsymbol{\mathcal{A}}^\sigma(\mathbf{R}) = -\frac{\sigma}{2}\left(1-\cos\theta(\mathbf{R})\right)\nabla\varphi(\mathbf{R}), \tag{163}$$

while its curl gives rise to the effective magnetic field. Both the vector potential given by the latter relation and the corresponding magnetic field are given by exactly the same expressions as the Berry connection and Berry curvature = emergent magnetic field from before. Moreover, analogously to the case of AB-effect, given the instantaneous eigenstates of H_0, $|\sigma\mathbf{e}_3\rangle_0$, we can construct the wavefunctions $|\mathbf{R}\sigma\mathbf{e}_3\rangle$ according to equation (36) where $\mathbf{A}$ should be replaced with $\boldsymbol{\mathcal{A}}^\sigma$, and this will give us the Berry connection identical to (163). It can be readily shown, be referring to the apparatus of section 2.3, that $|\mathbf{R}\sigma\mathbf{e}_3\rangle$ is the instantaneous eigenstate of the following Hamiltonian:

$$H^\sigma_{\text{eff}} = \frac{1}{2}\left(\mathbf{p}+\boldsymbol{\mathcal{A}}^\sigma\right)^2 - JM\sigma \tag{164}$$

We have thus explicitly shown that the Berry phase problem of Hamiltonian (154) can be also recast as an Aharonov-Bohm Berry phase problem of an effective ferromagnetic medium subject to a magnetic field, opposite for electrons of opposite spin. This field is identical to the emergent magnetic field, and gives rise to the topological Hall effect.

General case.
While a detailed analysis of the model Hamiltonian given by (154) is extremely insightful, the

realistic Hamiltonian of a solid which exhibits a spatially varying magnetic texture is more complex, in particular, it includes the crystal potential and spin-orbit interaction. From the viewpoint of the previous section, the Hamiltonian of our system can be written as $H = H(\mathbf{k}, \mathbf{R})$, where we take the $\mathbf{k}$ as the "local" Bloch vector at point $\mathbf{R}$, we exclude explicit dependence of the Hamiltonian on time (due to e.g. moving spin-texture), and we assume that the crystal potential does not depend on $\mathbf{R}$, so that the $\mathbf{R}$-dependence appears only due to magnetic texture. For each band the components of the connection $\mathcal{A}^n$ and curvature Ω^n forms in $(\mathbf{R}, \mathbf{k})$ space are given by Eqs. (131), (132) and (133).

Let us first consider the case without spin-orbit coupling. In this case the spin part and the orbital part of the wavefunctions are decoupled from each other. For a locally ferromagnetic crystal at point $\mathbf{R}$ without spin-orbit interaction $\Omega^n_{\mathbf{kk}} \equiv 0$. It is also clear that without spin-orbit coupling $\Omega^n_{\mathbf{kR}} = \partial_{\mathbf{R}}\mathcal{A}^n_{\mathbf{k}} - \partial_{\mathbf{k}}\mathcal{A}^n_{\mathbf{R}} \equiv 0$, since without spin-orbit $\mathcal{A}^n_{\mathbf{k}}$ does not depend on $\mathbf{R}$, and $\mathcal{A}^n_{\mathbf{R}}$ does not depend on $\mathbf{k}$. Thus, the only non-vanishing component of the curvature is the $\Omega^n_{\mathbf{RR}}$. In this case the only effect of the varying spin-texture is the spin-rotation of the wavefunctions for each band and at each $\mathbf{k}$ into the local spin-quantization axis, specified by $\mathbf{R}$. Since without spin-orbit coupling rotation of the spin-quantization axis does not lead to any changes in the spectrum or orbital parts of the wavefunctions, the problem can be effectively re-written in the spin-space, and the apparatus we developed before can be employed to the full extent. Namely, the $\Omega^n_{\mathbf{RR}}$ corresponds to the $\Omega^n_{\mathbf{R}}$ from before, and it is *exactly* given by Eq. (156).

Without spin-orbit interaction, the relation between the topological and ordinary Hall effects is essentially the same as that between the AHE and SHE (without spin-flip SOC): while one is the sum, the other is the difference of the intrinsic Hall (ordinary Hall in case of THE and OHE) effects for spin-up and spin-down electrons separately. Since the texture varies in space slowly, the electrons experience a sufficient amount of disorder-driven momentum-scattering on the scale of a texture, so that the relaxation approximation to the Boltzmann equation could be used, leading to the following expression for the topological Hall conductivity (for the texture in-plane and emergent field out of plane) [18]:

$$\sigma_{xy}^{\mathrm{THE}} \approx \sum_{\sigma} \sigma |B_e^{\sigma}| \sum_n \int_{\mathrm{BZ}} \tau_{\sigma n}^2 \left(\frac{(v^x_{\mathbf{k}n})^2}{m^{yy}_{\mathbf{k}n}} - \frac{v^x_{\mathbf{k}n} v^y_{\mathbf{k}n}}{m^{xy}_{\mathbf{k}n}} \right) \frac{\partial f_0(\varepsilon^{\sigma}_{\mathbf{k}n})}{\partial \varepsilon} \frac{d^3k}{(2\pi)^3}, \tag{165}$$

where $m_{\mathbf{k}n}$ is the mass tensor, and $\tau_{\sigma n}$ are the relaxation times. The ordinary Hall conductivity obtained within the same approximations in the orbital magnetic field $|B_e^{\sigma}|$ would read the same as the expression above, but without σ explicitly in the sum.

Mixed Berry curvature, Dzyaloshinskii-Moriya interaction and charging of skyrmions.

In the presence of spin-orbit coupling the $\Omega^n_{\mathbf{kk}}$ part of the Berry curvature tensor is different from zero, and this leads to the co-existence of the anomalous Hall and topological Hall contributions to the transverse conductivity, as seen experimentally. Another manifestation of the spin-orbit interaction is the fact that the mixed Berry curvature $\Omega^n_{\mathbf{kR}}$ is not anymore vanishing. A profound consequence of this fact is that the modified density of states in $(\mathbf{k}, \mathbf{R})$-space, given by (144), neglecting second and higher order terms, is given by [19]:

$$D^n(\mathbf{k}, \mathbf{R}) = \frac{1}{(2\pi)^d} \left(1 - \sum_i \Omega^n_{\mathbf{kR},ii} \right). \tag{166}$$

We can now write again the expression (151) for the free energy at a certain point $\mathbf{R}$, as we did

for deriving the orbital magnetization, but now with a different modified density of states:

$$F(\mathbf{R}) = -\frac{1}{\beta(2\pi)^d}\sum_n \int d\mathbf{k} \left(1 - \sum_i \Omega^n_{\mathbf{kR},ii}\right) \ln\left(1 + e^{-\beta(\mathcal{E}_{n\mathbf{kR}}-\mu)}\right), \tag{167}$$

where, as previously, $\mathcal{E}_{n\mathbf{kR}}$ is given by (136) and contains a correction $\delta\varepsilon_{n\mathbf{kR}}$ due to the texture in addition to the band energy. In the case of the constant magnetic field we were able to reduce the expression for $\delta\varepsilon_{n\mathbf{kR}}$ to $\mathbf{k}$-derivatives only. This cannot be done for the spin textures however, since the dependence of the wavefunctions on $\mathbf{R}$ is not only via the $\mathbf{k}$-vector:

$$\delta\varepsilon_{n\mathbf{kR}} = -\mathrm{Im}\,\langle \partial_{\mathbf{R}} u_{n\mathbf{kR}} \,|\, \varepsilon_{n\mathbf{kR}} - H_{\mathbf{kR}} \,|\, \partial_{\mathbf{k}} u_{n\mathbf{kR}}\rangle \tag{168}$$

We can now extract the contribution to the free energy which comes from the chirality of the magnetization. To first order in gradients of the magnetization, it reads:

$$\delta F(\mathbf{R}) = \frac{1}{(2\pi)^d}\sum_n \int d\mathbf{k} \left(f_{\mathbf{k}n}\,\delta\varepsilon_{n\mathbf{kR}} + \frac{\sum_i \Omega^n_{\mathbf{kR},ii}}{\beta} \ln\left(1 + e^{-\beta(\varepsilon_{n\mathbf{kR}}-\mu)}\right)\right), \tag{169}$$

where $f_{\mathbf{k}n}$ is the Fermi occupation function. The correction due to the chirality can be expanded in terms of the gradient of the magnetization as:

$$\delta F(\mathbf{R}) = D_{ij}(\mathbf{R})\,\hat{\mathbf{e}}_i \cdot \left(\hat{\mathbf{n}} \times \partial_{R_j}\hat{\mathbf{n}}\right), \tag{170}$$

where D_{ij} correspond to the *Dzyaloshinskii-Moriya interaction* (DMI) [20], which energetically favors the chirality of magnetization and which is discussed in detail by Stefan Blügel in his manuscript C4 of the current book. From (169) we can derive that

$$D_{ij}(\mathbf{R}) = \frac{1}{(2\pi)^d}\sum_n \int d\mathbf{k} \left(f_{\mathbf{k}n} A^{ij}_{n\mathbf{kR}} + \frac{B^{ij}_{n\mathbf{kR}}}{\beta} \ln\left(1 + e^{-\beta(\varepsilon_{n\mathbf{kR}}-\mu)}\right)\right), \tag{171}$$

with

$$A^{ij}_{n\mathbf{kR}} = - \mathbf{e}_i \cdot \mathbf{e}_\varphi\, \mathrm{Im}\,\langle \partial_\theta u_{n\mathbf{kR}} | \varepsilon_{n\mathbf{kR}} - H_{\mathbf{kR}} | \partial_{k_j} u_{n\mathbf{kR}}\rangle \tag{172}$$

$$- (\mathbf{e}_i \cdot \mathbf{e}_\theta / \sin\theta)\, \mathrm{Im}\,\langle \partial_\varphi u_{n\mathbf{kR}} | \varepsilon_{n\mathbf{kR}} - H_{\mathbf{kR}} | \partial_{k_j} u_{n\mathbf{kR}}\rangle, \tag{173}$$

and

$$B^{ij}_{n\mathbf{kR}} = -2\mathbf{e}_i \cdot \left[\mathbf{e}_\varphi\, \mathrm{Im}\,\langle \partial_\theta u_{n\mathbf{kR}} | \partial_{k_j} u_{n\mathbf{kR}}\rangle - (\mathbf{e}_\theta / \sin\theta)\, \mathrm{Im}\,\langle \partial_\varphi u_{n\mathbf{kR}} | \partial_{k_j} u_{n\mathbf{kR}}\rangle\right], \tag{174}$$

where $\mathbf{e}_i$ are the cartesian unit vectors, while $\mathbf{e}_\theta = \partial\mathbf{n}/\partial\theta$ and $\mathbf{e}_\varphi = (1/\sin\theta)\partial\mathbf{n}/\partial\varphi$ are the unit vectors on the sphere, Fig. 1. At zero temperature the DMI can be written as:

$$D_{ij}(\mathbf{R}) = \frac{1}{(2\pi)^d}\sum_n \int d\mathbf{k}\, f_{\mathbf{k}n} \left[A^{ij}_{n\mathbf{kR}} - (\varepsilon_{n\mathbf{kR}} - \mu) B^{ij}_{n\mathbf{kR}}\right]. \tag{175}$$

This expression has a one-to-one resemblance to the formula for the orbital magnetization in ferromagnets, Eq. (153), which leads to naming D_{ij} the *DMI spiralization* [20]. The quantity $A^{ij}_{n\mathbf{kR}}$ is called the *twist torque moment* of state n and it corresponds to the local orbital moment of a wavepacket in case of the orbital magnetiation. Obviously, the $B^{ij}_{n\mathbf{kR}}$ presents the correction to the DMI spiralization due to the mixed Berry curvature, playing the same role the $\mathbf{k}$-space

Berry curvature plays for the orbital magnetization [20]. Besides fundamental importance, the Berry phase expression (175) can be used to compute the DMI spiralization from the electronic structure of the collinear system with the magnetization pointing along the direction which corresponds to $\mathbf{R}$. This presents a great simplification as compared to current approaches, used to calculate the DMI in solids by explicitly incorporating the chirality of the magnetization into the calculations. For example, in left handed crystal structure of MnSi the $D_{ij} = -D\,\delta_{ij}$ and the D can be calculated to be -4.1 meV per Å and per 8-atom unit cell, in good agreement to experiment [19].
Analogously to the case of the change in the Fermi volume due to modified density of states we calculated before, we can also consider the change in the charge density due to chirality, which comes from the change in the energy of the states and modified density of states:

$$\delta\rho(\mathbf{R}) = \sum_n \frac{1}{(2\pi)^d} \int d\mathbf{k} \left(\frac{\partial f_{n\mathbf{k}}}{\partial \varepsilon}\, \delta\varepsilon_{n\mathbf{kR}} - f_{n\mathbf{k}} \sum_i \Omega^n_{\mathbf{kR},ii}. \right) \tag{176}$$

The change of the charge density consists of two terms, of which the first one is the Fermi surface and the second one is the Fermi see contribution. Formally, the Fermi see term corresponds to expression (77) for the derivative of the electric polarization of an insulator with respect to $\lambda = \mathbf{R}$. In the latter case the integral of this derivative over λ would give the change of polarization upon varying λ. Similarly, the integral of (176) over the $\mathbf{R}$-texture gives the additional electric charge $\delta\rho$ of the skyrmion due to the Berry phase effect. There are two important differences between the two situations, however. The first is that in insulating skyrmions the integral of the mixed Berry curvature over $\mathbf{k}$ and $\mathbf{R}$ spaces vanishes [19]. The second one is that in metals there is also a Fermi surface contribution to the charge, which also vanishes for insulating textures. Thus, to linear order, the charge of insulating skyrmions is zero and the charging effects in this case occur due to higher-order terms, which can be recast in terms related to the Chern numbers. In metals, on the other hand, the charge evaluated according to (176) will be screened by conduction electrons and its value will be significantly reduced. So, in MnSi for example, the computed unscreened value of $\delta\rho$ is about 0.25 electrons, while the screened charge is reduced by orders of magnitude with the distribution of $\delta\rho$ over the skyrmion still significant [19].

References

[1] M. V. Berry, Proc. R. Soc. Lond. A **392**, 45 (1984)

[2] A. Bohm, A. Mostafazadeh, H. Koizumi, Q. Niu, and J. Zwanziger, *The Geometric Phase in Quantum Systems*, Springer-Verlag Berlin Heidelberg (2003)

[3] M. Nakahara, *Geometry, Topology and Physics*, IoP Publishing Ltd (2003)

[4] R. Resta and D. Vanderbilt, *Theory of Polarization: A Modern Approach* in *Physics of Ferroelectrics: a Modern Perspective*, Springer (2007)

[5] B. A. Bernevig and T. L. Hughes, *Topological Insulators and Topological Superconductors*, Princeton University Press (2013).

[6] D. Xiao, M.-C. Chang and Q. Niu, Rev. Mod. Phys. **82**, 1959 (2010)

[7] T. Fujita, M. B. A. Jalil, S. G. Tan, and S. Murakami, J. Appl. Phys. **110**, 121301 (2011)

[8] R. Resta, Ferroelectrics **136**, 51 (1992)

[9] R. D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993)

[10] J. Bellissard, arXiv:cond-mat/9504030

[11] M.-Y. Lee, M.-C. Chang, and T.-M. Hong, Phys. Rev. B **57**, 11895 (1998)

[12] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982)

[13] N. Nagaosa, J. Sinova, S. Onoda, A. H.MacDonald, and N. P. Ong, Rev. Mod. Phys. **82**, 1539 (2010)

[14] F. D. M. Haldane, Phys. Rev. Lett. **61**, 2015 (1988)

[15] H. Zhang, F. Freimuth, G. Bihlmayer, M. Ležaić, S. Blügel and Y. Mokrousov, Phys. Rev. B **87**, 205132 (2013)

[16] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005)

[17] G. Sundaram and Q. Niu, Phys. Rev. B **59**, 14915 (1999)

[18] R. Ritz, M. Halder, C. Franz, A. Bauer, M. Wagner, R. Bamler, A. Rosch and C. Pfleiderer, Phys. Rev. B **87**, 134424 (2013)

[19] F. Freimuth, R. Bamler, Y. Mokrousov and A. Rosch, Phys. Rev. B **88**, 214409 (2013)

[20] F. Freimuth, S. Blügel, and Y. Mokrousov, arXiv:1308.5983 (2013), J. Phys.: Cond. Mat. (2014)

A 7 Time-dependent density-functional theory [1]

N. Helbig
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Time-dependent density functional theory (TDDFT) extends the density functional formalism to include time-dependent external potentials which frequently occur for example if the system is exposed to an external laser field. TDDFT allows for the calculation of the the excitation energies of the system which can be compared to experimental photo-absorption spectra. As in static DFT there are approximations involved and the results depend on the quality of these approximations. In this chapter we give a short introduction into the subject starting from the Runge-Gross theorem, the time-dependent equivalent of the Hohenberg-Kohn theorem. We introduce the time-dependent Kohn-Sham (KS) system and discuss the linear response scheme which is used in most applications today. We emphasize, however, that TDDFT is not restricted to small perturbations, the fundamental theory covers almost any kind of time-dependent external potential of interest. We conclude the chapter with two examples for very basic applications. In order to keep the notation simple and accessible for non-experts, we only consider the spin-independent version of TDDFT. The theory can be extended to include the spin degrees of freedom in the same way as static DFT. Obviously, this chapter cannot provide a complete introduction into TDDFT or discuss the more subtle points of the theory. A more detailed and extended overview of TDDFT can be found in [1, 2].
The behavior of a many-body system exposed to a time-dependent external potential is described by the time-dependent Schrödinger equation

$$\hat{H}(t)\Psi(\mathbf{r}_1,...\mathbf{r}_N,t) = i\partial_t\Psi(\mathbf{r}_1,...\mathbf{r}_N,t) \tag{1}$$

with (we are using atomic units throughout this chapter unless explicitly stated otherwise)

$$\hat{H}(t) = \sum_{j=1}^{N}\left(-\frac{\nabla_j^2}{2} + V(\mathbf{r}_j,t)\right) + \frac{1}{2}\sum_{\substack{j,k=1\\ j\neq k}}^{N}\frac{1}{|\mathbf{r}_j-\mathbf{r}_k|}. \tag{2}$$

As the Schrödinger equation is a first order differential equation in time the wave function $\Psi(\mathbf{r}_1,...\mathbf{r}_N,t)$ is fully determined once we specify an initial state $\Psi(\mathbf{r}_1,...\mathbf{r}_N,t_0)$ at some initial time t_0. Hamiltonians of the form (2) allow, for example, for the treatment of external laser fields which are commonly included in the dipole approximation as $v(\mathbf{r},t) = \mathbf{r}\mathbf{E}(t)$ with $\mathbf{E}(t)$ being the electric field of the laser.
The foundation of TDDFT lies in the Runge-Gross theorem [3], a time-dependent version of the Hohenberg-Kohn theorem, which establishes a one-to-one correspondence between the time-dependent potential $V(\mathbf{r},t)$ and the time-dependent density

$$n(\mathbf{r},t) = N\int d^3r_2...d^3r_N|\Psi(\mathbf{r},\mathbf{r}_2...\mathbf{r}_N,t)|^2 \tag{3}$$

for the evolution of a system starting from a given initial state $\Psi(\mathbf{r}_1,...\mathbf{r}_N,t_0)$ at t_0 as long as the potential $V(\mathbf{r},t)$ is Taylor-expandable around the initial time t_0. Unfortunately, this excludes all processes where a field is switched on adiabatically starting at $t\to-\infty$. The extension of the theorem to non-analytic potentials is an ongoing effort in the TDDFT community at present, see for example [4]. Here, we review the original proof by Runge and Gross before introducing the time-dependent Kohn-Sham system and discussing the approximations that need to be made in time-dependent DFT which go beyond those of its static counterpart.

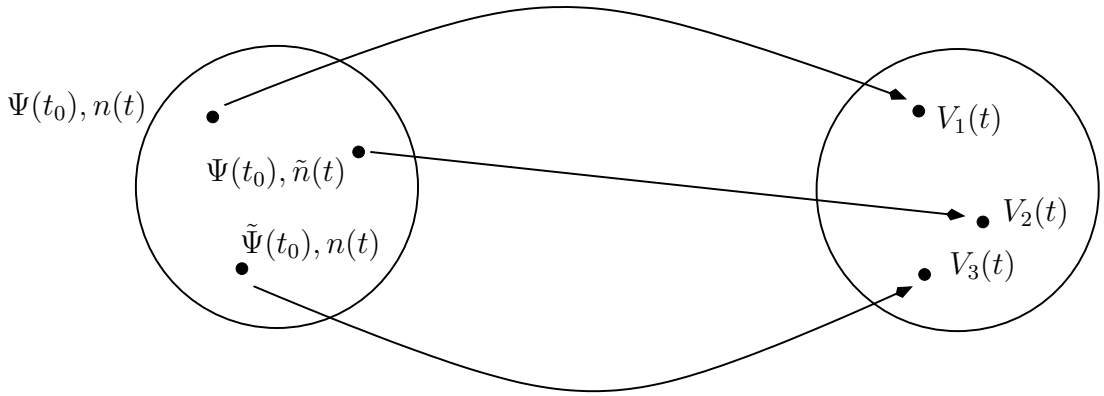


Fig. 1: *One-to-one correspondence between elements of the set of pairs of initial states $\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N, t_0)$ and densities $n(\mathbf{r}, t)$ (on the left) and elements of the set Taylor-expandable potentials $V(\mathbf{r}, t)$ (on the right). Changing only the initial state already corresponds to a different potential even if the density remains unchanged. The three potentials V_1, V_2 and V_3 differ by more than a purely time-dependent function. (All spatial variables have been suppressed for clarity in the picture.)*

2 Formalism

The goal of TDDFT is to avoid the calculation of the time-dependent wave function of interacting electrons, $\Psi(\mathbf{r}, \mathbf{r}_2 ... \mathbf{r}_N, t)$, and replace it with the calculation of only the density $n(\mathbf{r}, t)$. Hence, we first need to prove that this density uniquely determines all properties of the system which implies that it uniquely determines the potential $V(\mathbf{r}, t)$ in Eq. (2). In analogy to static DFT we would then like to introduce an effective system of non-interacting electrons with the same time-dependent density as the many-body system, in other words, we would like to construct a time-dependent Kohn-Sham (KS) system. All properties of the interacting many-body system are then calculated as functionals of $n(\mathbf{r}, t)$ which is determined from the KS system. The task of solving the time-dependent Schrödinger equation for the interacting system is hence replaced by the task of finding the functional dependence of an observable of interest on the time-dependent density.

2.1 Runge-Gross Theorem

The Runge-Gross theorem [3] states that two densities $n(\mathbf{r}, t)$ and $n'(\mathbf{r}, t)$ evolving from a common initial state $\Psi(\mathbf{r}_1 ... \mathbf{r}_N, t_0)$ under the influence of two potentials $V(\mathbf{r}, t)$ and $V'(\mathbf{r}, t)$ which are both Taylor-expandable about the initial time t_0, eventually differ if the potentials differ by more than a purely time-dependent function, i.e. if $V(\mathbf{r}, t) - V'(\mathbf{r}, t) \neq c(t)$. In other words, the initial state and the time-dependent density *together* determine the potential up to a purely time-dependent function. Therefore, using two different initial states, $\Psi(\mathbf{r}_1 ... \mathbf{r}_N, t_0)$ and $\tilde{\Psi}(\mathbf{r}_1 ... \mathbf{r}_N, t_0)$, with the same initial density $n(\mathbf{r}, t_0)$, and require the density to remain also for $t > t_0$ we need to use a different potential. A graphical representation of the Runge-Gross theorem is given in Fig. 1 where it is understood that all three potentials differ by more than a purely time-dependent function. The opposite direction, that the potential determines the density uniquely for a given initial state, is an obvious consequence of the unique solution of the time-dependent Schrödinger equation and the definition of the density.

Since the two potentials $V(\mathbf{r},t)$ and $V'(\mathbf{r},t)$ are required to be analytical around t_0 we can expand them as

$$V(\mathbf{r},t) = \sum_{k=0}^{\infty} \frac{1}{k!} V_k(\mathbf{r})(t-t_0)^k, \tag{4}$$

$$V'(\mathbf{r},t) = \sum_{k=0}^{\infty} \frac{1}{k!} V'_k(\mathbf{r})(t-t_0)^k. \tag{5}$$

The fact that the two potentials differ by more than a purely time-dependent function then implies that there exists a k for which

$$V_k(\mathbf{r}) - V'_k(\mathbf{r}) = \frac{\partial^k}{\partial t^k} \left[V(\mathbf{r},t) - V'(\mathbf{r},t)\right]_{t=t_0} \neq \text{const.} \tag{6}$$

In other words, there exists an order k in the Taylor expansion with a difference in the expansion coefficients of V and V' that is space dependent.
Following the original proof [3] we first establish a one-to-one correspondence between the current density

$$\mathbf{j}(\mathbf{r},t) = \frac{1}{2i} \int d^3r_2 \ldots \int d^3r_N \left[\Psi^*(\mathbf{r},\mathbf{r}_2 ... \mathbf{r}_N,t)\nabla\Psi(\mathbf{r},\mathbf{r}_2 ... \mathbf{r}_N,t) - \left(\nabla\Psi^*(\mathbf{r},\mathbf{r}_2 ... \mathbf{r}_N,t)\right)\Psi(\mathbf{r},\mathbf{r}_2 ... \mathbf{r}_N,t)\right] \tag{7}$$

and the potential $V(\mathbf{r},t)$ before using the continuity equation to go from the current density to the density in a second step.

Step 1: The equations of motion for the current densities $\mathbf{j}$ and $\mathbf{j}'$ read as

$$\frac{\partial}{\partial t}\mathbf{j}(\mathbf{r},t) = \frac{\partial}{\partial t}\langle\Psi(t)|\hat{\mathbf{j}}(\mathbf{r})|\Psi(t)\rangle = -i\langle\Psi(t)|[\hat{\mathbf{j}}(\mathbf{r}),\hat{H}(t)]|\Psi(t)\rangle, \tag{8}$$

$$\frac{\partial}{\partial t}\mathbf{j}'(\mathbf{r},t) = \frac{\partial}{\partial t}\langle\Psi'(t)|\hat{\mathbf{j}}(\mathbf{r})|\Psi'(t)\rangle = -i\langle\Psi'(t)|[\hat{\mathbf{j}}(\mathbf{r}),\hat{H}'(t)]|\Psi'(t)\rangle. \tag{9}$$

The time-derivative of the operator which appears in the general equation is zero here since $\hat{\mathbf{j}}$ is not explicitly time-dependent. Taking the difference of Eqs. (8) and (9) at $t = t_0$ we obtain

$$\frac{\partial}{\partial t}\left[\mathbf{j}(\mathbf{r},t) - \mathbf{j}'(\mathbf{r},t)\right]\Big|_{t=t_0} = -i\langle\Psi(t_0)|[\hat{\mathbf{j}}(\mathbf{r}),\hat{H}(t_0) - \hat{H}'(t_0)]|\Psi(t_0)\rangle. \tag{10}$$

The only difference between the two Hamiltonians $\hat{H}$ and $\hat{H}'$ is in the potentials V and V', which allows us to write

$$\begin{aligned}\frac{\partial}{\partial t}\left[\mathbf{j}(\mathbf{r},t) - \mathbf{j}'(\mathbf{r},t)\right]\Big|_{t=t_0} &= -i\langle\Psi(t_0)|[\hat{\mathbf{j}}(\mathbf{r}),V(\mathbf{r},t_0) - V'(\mathbf{r},t_0)]|\Psi(t_0)\rangle \\ &= -n(\mathbf{r},t_0)\nabla\left(V(\mathbf{r},t_0) - V'(\mathbf{r},t_0)\right).\end{aligned} \tag{11}$$

If Eq. (6) holds for $k = 0$, i.e. the difference of the two potentials at $t = t_0$ is not constant over space, then the right-hand-side in Eq. (11) does not vanish and $\mathbf{j}$ and $\mathbf{j}'$ become different infinitesimally later than $t = t_0$. If the smallest k for which Eq. (6) holds is

non-zero we calculate the corresponding $(k+1)$th time derivative of the current which involves k nested commutators of $\hat{j}$ and $\hat{H}$ (or the corresponding primed quantities) and leads to

$$\frac{\partial^{k+1}}{\partial t^{k+1}} \left[\mathbf{j}(\mathbf{r},t) - \mathbf{j}'(\mathbf{r},t)\right]\Big|_{t=t_0} = -n(\mathbf{r},t_0)\nabla w_k(\mathbf{r}) \neq 0 \tag{12}$$

which is non-zero due to

$$w_k(\mathbf{r}) = \frac{\partial^k}{\partial t^k}\left[V(\mathbf{r},t) - V'(\mathbf{r},t)\right]\Big|_{t=t_0} = V_k(\mathbf{r}) - V_k'(\mathbf{r}) \neq \text{const.} \tag{13}$$

being different from a constant. Again, due to the right-hand-side in Eq. (12) being non-zero the two currents $\mathbf{j}$ and $\mathbf{j}'$ differ infinitesimally later than the initial time t_0. Therefore, we have succeeded in proving a one-to-one correspondence between the potential $V(\mathbf{r},t)$ and the current density $\mathbf{j}(\mathbf{r},t)$ provided that we use the same initial state for both the primed and the unprimed system.

Step 2: We now use the continuity equation

$$\frac{\partial n(\mathbf{r},t)}{\partial t} = -\nabla \mathbf{j}(\mathbf{r},t) \tag{14}$$

to prove the one-to-one correspondence between the potential and the density. Applying $(k+1)$ time derivatives to the continuity equation for both $n(\mathbf{r},t)$ and $n'(\mathbf{r},t)$ and using Eq. (12) we obtain

$$\begin{aligned}\frac{\partial^{k+2}}{\partial t^{k+2}}\left[n(\mathbf{r},t) - n'(\mathbf{r},t)\right]\Big|_{t=t_0} &= -\nabla \frac{\partial^{k+1}}{\partial t^{k+1}}\left[\mathbf{j}(\mathbf{r},t) - \mathbf{j}'(\mathbf{r},t)\right]\Big|_{t=t_0} \\ &= \nabla\left[n(\mathbf{r},t_0)\nabla w_k(\mathbf{r})\right]. \end{aligned} \tag{15}$$

Without the divergence operator in the last line of Eq. (15) our task would be complete since the term in the square brackets is non-zero. However, now we have to ensure in addition that the square bracket is not some non-zero constant. To this end we consider the integral

$$\begin{aligned}\int d^3r\, n(\mathbf{r},t_0)\left[\nabla w_k(\mathbf{r})\right]^2 &= -\int d^3r\, w_k(\mathbf{r})\nabla\left[n(\mathbf{r},t_0)\nabla w_k(\mathbf{r})\right] \\ &\quad + \oint d\mathbf{S}\left[n(\mathbf{r},t_0) w_k(\mathbf{r})\nabla w_k(\mathbf{r})\right]. \end{aligned} \tag{16}$$

For finite systems, where the potential arises from a finite normalizable charge distribution, one can show that the potential decays at least as fast as $1/r$. Therefore, by choosing the surface to be at $r \to \infty$ the surface term vanishes. The integrand on the left of Eq. (16) is strictly non-negative and only zero on isolated points in space where either the density or $w_k(\mathbf{r})$ are zero. Hence, the whole integral cannot vanish which implies that the integrand in the first term on the right-hand-side cannot be zero at all points in space. Hence, the divergence appearing in Eq. (15) is non-zero. Using the same argument as before for the current density we can conclude that the two densities $n(\mathbf{r},t)$ and $n'(\mathbf{r},t)$ differ infinitesimally later than the initial time t_0. This concludes the proof of the one-to-one correspondence between the potential $V(\mathbf{r},t)$ and the density $n(\mathbf{r},t)$ for a given initial state $\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N, t_0)$.

As we have seen, the original Runge-Gross proof relies on the surface integral in Eq. (16) to vanish. While this is satisfied for finite systems the situation is more difficult for periodic systems. If one can find a periodicity, possibly with a larger unit cell than the periodic system without the time-dependent part of the external potential, the integrand in the surface term is identical on two opposite sites of this (larger) unit cell such that the contributions to the integral cancel. This excludes the case where one switches on a uniform electric field [5]. However, even if the surface term does not vanish the first step of the proof remains valid. Therefore, one can use the current density as the basic variable which is included in time-dependent current density functional theory which, however, additionally allows for external vector potentials [6].

2.2 Time-dependent Kohn-Sham equations

In analogy to static DFT we would like to construct an effective system of non-interacting electrons which reproduces the density of the interacting electrons. In the whole proof of the Runge-Gross theorem it is not specified whether the Hamiltonians $\hat{H}$ and $\hat{H}'$ contain any electron-electron interaction or not. Therefore, the proof simultaneously shows that the external potential in an interacting system is uniquely given by the density and the initial state and that the effective potential of a non-interacting system is a unique potential of the same time-dependent density and the non-interacting initial state. However, just as the Hohenberg-Kohn theorem [9] in static DFT, the Runge-Gross theorem does not provide insight into the existence of a non-interacting system which has a specific time-dependent density. This proof was provided by van Leeuwen [7] who showed that for a given many-body initial state and a given time-dependent density there exists an effective potential in a non-interacting system which yields the same density at all times provided one can find a non-interacting initial state, i.e. a Slater determinant, which yields both the correct density as well as the correct first time derivative of the density at the initial time. In case the first time derivative at the initial time is zero one can always construct a Slater determinant for a given initial density by the Harriman construction [8].
The time-dependent Kohn-Sham (KS) system is constructed such that it reproduces the density of the many-body system at all times t starting from the initial time t_0. The time evolution of the Kohn-Sham orbitals $\varphi_j(\mathbf{r},t)$ then follows from the single-particle Schrödinger equation

$$i\frac{\partial}{\partial t}\varphi_j(\mathbf{r},t) = \left[-\frac{\nabla^2}{2} + V_s(\mathbf{r},t)\right]\varphi_j(\mathbf{r},t), \tag{17}$$

where the KS potential $V_s(\mathbf{r},t)$ is the effective single-particle potential. The Runge-Gross proof then states that $V_s(\mathbf{r},t)$ is unique for a given initial state. For the non-interacting KS system the initial state is a Slater determinant built from the KS orbitals which we denote as $\Phi(\mathbf{r}_1 \ldots \mathbf{r}_N, t_0)$. Due to the many-body system being uniquely described by the density $n(\mathbf{r},t)$ and the many-body initial state $\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N, t_0)$ the potential $V_s(\mathbf{r},t)$ is a functional of the time-dependent density $n(\mathbf{r},t)$, the initial many-body state $\Psi(\mathbf{r}_1 \ldots \mathbf{r}_N, t_0)$ and the initial KS Slater determinant $\Phi(\mathbf{r}_1 \ldots \mathbf{r}_N, t_0)$, i.e.

$$V_s(\mathbf{r},t) = V_s\left[n, \Psi(t_0), \Phi(t_0)\right](\mathbf{r},t), \tag{18}$$

where we suppressed all arguments except the initial time to simplify the notation. If we choose the initial states to be the ground states of the many-body system and the KS system we can employ the Hohenberg-Kohn [9] theorem which ensures that these two states are functionals

of the ground-state density which is identical to the density at $t = t_0$. Hence, the effective potential $V_s(\mathbf{r}, t)$ becomes a functional of the time-dependent density alone. In all other cases, the functional has to depend on both initial states.
We choose to write the KS potential, in analogy to static DFT, as

$$V_s(\mathbf{r}, t) = V_0(\mathbf{r}, t) + V_H(\mathbf{r}, t) + V_{xc}(\mathbf{r}, t) \tag{19}$$

with $V_0(\mathbf{r}, t)$ being the external potential and the Hartree potential being given as

$$V_H(\mathbf{r}, t) = \int d^3r' \frac{n(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|}. \tag{20}$$

The remaining unknown part of V_s we denote as the exchange-correlation potential V_{xc}. While in static DFT the form of the KS potential is determined by the condition that the interacting and the KS systems have the same ground-state density, i.e. their energy is minimized by the same density, there is no such condition in time-dependent DFT. Eq. (19) can, therefore, be regarded as a definition of the exchange-correlation potential $V_{xc}(\mathbf{r}, t)$. Due to the effective potential $V_s(\mathbf{r}, t)$ being a functional of the time-dependent density and the two initial states, also $V_{xc}(\mathbf{r}, t)$ is a functional of all three quantities unless we start in the ground states at $t = t_0$.
The KS scheme discussed up to this point is in principle exact, if we had the exact functional V_{xc} and the correct initial states we would obtain the correct time-dependent density from the propagation of the KS equations. However, in practise those things are unknown leading to the calculations being performed in several steps each connected to its own approximations

- Perform a static DFT calculation to obtain the ground-state of the system as an initial state for the subsequent time propagation. This requires an approximation to the exchange-correlation energy from which the static exchange-correlation potential can be derived. The result is a set of occupied KS orbitals.

- Perform a time-dependent DFT calculation to propagate the occupied KS orbitals. This requires an approximation to the time-dependent exchange-correlation potential $V_{xc}(\mathbf{r}, t)$. The fact that this potential not only depends on the time-dependent density but also on the initial KS state and the initial many-body state of course complicates the search for appropriate approximations. An additional complication lies in the fact that in order to calculate the potential at a given time t not only the density at that time enters but the whole history of the evolution of the density since the initial time t_0 is necessary to obtain a unique answer. In order not to violate causality, the potential $V_{xc}(\mathbf{r}, t)$ obviously cannot depend on densities at times later than t.

- For practical calculations we need an algorithm for the numerical time propagation which does not violate time-reversal symmetry if the Hamiltonian has that symmetry. For a discussion of possible propagation schemes see Ref. [12].

2.3 Approximations to $V_{xc}(\mathbf{r}, t)$

As discussed in the previous section, the exchange-correlation potential $V_{xc}(\mathbf{r}, t)$, i.e. the part of the KS potential that needs to be approximated, is generally a functional of the time-dependent density $n(\mathbf{r}, t)$, the initial many-body state $\Psi(\mathbf{r}_1, \ldots \mathbf{r}_N, t_0)$ and the initial KS Slater determinant $\Phi(\mathbf{r}_1, \ldots \mathbf{r}_N, t_0)$. If the many-body and the KS systems are in their respective ground states

at t_0 we can employ the Hohenberg-Kohn theorem [9] which shows that $\Psi(\mathbf{r}_1, \dots \mathbf{r}_N, t_0)$ and $\Phi(\mathbf{r}_1, \dots \mathbf{r}_N, t_0)$ are functionals of the density $n(\mathbf{r}, t_0)$. Hence, the KS potential V_s and consequently also V_{xc} are functionals of the density alone. However, the dependence on the history of the propagation remains. In other words, the potential at time t depends on all densities in the time interval $[t_0, t]$ from the initial time to the time for which the potential is calculated. The search for truly history-dependent approximations is ongoing [10, 11]. In most applications it is simply neglected and one uses a so-called adiabatic approximation which we discuss in the following.

Adiabatic approximations

The simplest way to construct an approximation for the time-dependent exchange-correlation potential uses an approximation to the exchange-correlation potential from static DFT and evaluates it at the time-dependent density, i.e.

$$V_{xc}(\mathbf{r}, t) = \left. \frac{\delta E_{xc}^{\text{static}}}{\delta n(\mathbf{r})} \right|_{n(\mathbf{r},t)}. \tag{21}$$

For example, using the local density approximation to the exchange energy

$$E_x^{\text{LDA}} = -\frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3} \int d^3r \, n(\mathbf{r})^{4/3} \tag{22}$$

the adiabatic LDA (ALDA) approximation to the time-dependent exchange potential reads

$$V_x^{\text{ALDA}}(\mathbf{r}, t) = -\left(\frac{3}{\pi} \right)^{1/3} n(\mathbf{r}, t)^{1/3}. \tag{23}$$

As we can see, the potential at time t only depends on the density at this time, neglecting all dependence on the history of the time evolution. This property is shared by all adiabatic functionals, they do not contain any memory of the time evolution. Also, since the approximation is borrowed from static DFT, there is no dependence on the initial states in adiabatic approximations. Nevertheless, the large majority of calculations in TDDFT are performed using the adiabatic local density approximation to $V_{xc}(\mathbf{r}, t)$.

2.4 Time-dependent observables

In quantum mechanics observables are calculated by evaluating the appropriate operator with respect to the wave function, i.e. at time t we evaluate

$$\langle \hat{O} \rangle = \langle \Psi(t) | \hat{O} | \Psi(t) \rangle. \tag{24}$$

Due to the Runge-Gross theorem the wave function $\Psi(t)$ is a functional of the time-dependent density and the initial state. Hence, if we had the exact exchange-correlation potential we could propagate the KS equations exactly and obtain the correct time-dependent density of the many-body system. If we knew in addition the functional dependence of the many-body wave function $\Psi(t)$ on the density we could calculate all observables that we might ever become interested in. Unfortunately, both the exact functional for the potential and for the wave function are generally unknown. Therefore, our results will almost always be approximate. In addition,

due to $\Psi[n, \Psi(t_0)](t)$ being unknown, we have to separately find the functional dependence on the density and the initial state of each observable that we are interested in. Depending on the observable this can be trivial or nearly impossible. One important example for an observable with a trivial dependence on the density is the dipole moment $\mathbf{d}(t)$ of the system which is given as

$$\mathbf{d}(t) = \int d^3r \, \mathbf{r} n(\mathbf{r}, t). \tag{25}$$

For the non-trivial cases one often uses the expectation value of the operator evaluated with the KS Slater determinant as an approximation. Since the KS Slater determinant is not supposed to be considered an approximation to the interacting wave function the quality of this approximate functional varies from observable to observable.

3 Linear Response

Many applications of time-dependent DFT rely on a situation where a system described by a static Hamiltonian $\hat{H}_0$ is perturbed with a time-dependent potential $\delta V_0(\mathbf{r}, t)$. In addition one only considers cases in which the system is in an eigenstate of $\hat{H}_0$, most often the ground state, when the perturbation is applied at time t', i.e. $\delta V_0(\mathbf{r}, t) = 0$ for $t < t'$. If the perturbation is sufficiently small the properties of the system at $t > t'$ can be calculated from first order perturbation theory, i.e. in linear response. This situation corresponds to a typical photo-electron experiment where one uses a laser field to determine the excitation energies of the system. The laser field has to be sufficiently weak in order for the excitation energies to remain unchanged by the laser.
As one maps the interacting system onto the KS system, the small change in the density due to the external perturbation can be obtained from the response of the KS system to the same external perturbation. Of course, due to the change in the density not only the external potential in the KS potential changes but also the Hartree and exchange-correlation contributions.

3.1 Response functions

In the following we very briefly discuss the basics of response functions mostly to introduce the notation. We assume that the system is originally in its ground state before the perturbation. A detailed introduction into response functions can be found in many books on many-body physics, e.g. [13].
The linear response functions of the interacting and the KS systems are defined as

$$\chi(\mathbf{r}, \mathbf{r}', t - t') = \left. \frac{\delta n(\mathbf{r}, t)}{\delta V_0(\mathbf{r}', t')} \right|_{n_0}, \tag{26}$$

$$\chi_s(\mathbf{r}, \mathbf{r}', t - t') = \left. \frac{\delta n(\mathbf{r}, t)}{\delta V_s(\mathbf{r}', t')} \right|_{n_0}, \tag{27}$$

where the derivatives are evaluated at the ground state density n_0. Both response functions depend only on the time that passes between the application of the perturbation at t' and the measurement of the change in the density at time t because the system is originally in its ground state Ψ_0 and Φ_0 for the many-body and the KS systems, respectively. Hence, the wave function only picks up an irrelevant total phase leaving the density unchanged until the perturbation is

applied. Applying a Fourier transformation with respect to $t - t'$ we obtain the corresponding response functions in frequency space which are given in their *Lehmann representation* as

$$\chi(\mathbf{r},\mathbf{r}',\omega) = \sum_J \frac{\langle\Psi_0|\hat{n}(\mathbf{r})|\Psi_J\rangle\langle\Psi_J|\hat{n}(\mathbf{r})|\Psi_0\rangle}{\omega-(E_0-E_J)+i0^+} - \frac{\langle\Psi_0|\hat{n}(\mathbf{r})|\Psi_J\rangle\langle\Psi_J|\hat{n}(\mathbf{r})|\Psi_0\rangle}{\omega+(E_0-E_J)+i0^+}, \quad (28)$$

$$\begin{aligned}\chi_s(\mathbf{r},\mathbf{r}',\omega) &= \sum_J \frac{\langle\Phi_0|\hat{n}(\mathbf{r})|\Phi_J\rangle\langle\Phi_J|\hat{n}(\mathbf{r})|\Phi_0\rangle}{\omega-(\mathcal{E}_0-\mathcal{E}_J)+i0^+} - \frac{\langle\Phi_0|\hat{n}(\mathbf{r})|\Phi_J\rangle\langle\Phi_J|\hat{n}(\mathbf{r})|\Phi_0\rangle}{\omega+(\mathcal{E}_0-\mathcal{E}_J)+i0^+} \\ &= \lim_{\eta\to 0^+}\sum_{j,k}(f_k-f_j)\frac{\varphi_k^*(\mathbf{r})\varphi_j(\mathbf{r})\varphi_j^*(\mathbf{r}')\varphi_k(\mathbf{r}')}{\omega-(\epsilon_j-\epsilon_k)+i\eta}, \quad (29)\end{aligned}$$

where E_J, $\mathcal{E}_J$ and ϵ_j denote the many-body eigenenergies, the eigenenergies of the KS Slater determinants and the KS energy eigenvalues, respectively. The density operator is written as $\hat{n}(\mathbf{r})$ and the occupations of the KS orbitals $\varphi_j(\mathbf{r})$ before the perturbation are given by f_j. As we can see, the response functions have a pole whenever ω equals the difference between two energy eigenvalues. Therefore, the many-body response function has its poles at the many-body excitation energies while the KS response function has poles at the difference of KS energy eigenvalues. As we will see in the following the two do not coincide showing that the KS energy differences should not be taken as excitation energies of a many-body system. We also note that the calculation of the response function involves an infinite sum over excited many-body states or unoccupied KS orbitals which hinders their numerical evaluation.

3.2 Dyson equation

Starting from the interacting response function we can use the fact that the KS system reproduces the change in the density of the interacting system. Therefore, this change can be calculated from the change in the KS potential leading to

$$\begin{aligned}\chi(\mathbf{r},\mathbf{r}',t-t') &= \int d\tau\int d^3x\,\frac{\delta n(\mathbf{r},t)}{\delta V_s(\mathbf{x},\tau)}\frac{\delta V_s(\mathbf{x},\tau)}{\delta V_0(\mathbf{r}',t')} \\ &= \int d\tau\int d^3x\,\chi_s(\mathbf{r},\mathbf{x},t-\tau) \\ &\quad\times\left[\delta(\mathbf{x}-\mathbf{r}')\delta(\tau-t')+\frac{\delta V_H(\mathbf{x},\tau)}{\delta V_0(\mathbf{r}',t')}+\frac{\delta V_{xc}(\mathbf{x},\tau)}{\delta V_0(\mathbf{r}',t')}\right], \quad (30)\end{aligned}$$

where we used the decomposition of V_s, Eq. (19). Employing a further chain rule to evaluate the remaining derivatives with respect to V_0 we obtain

$$\begin{aligned}\chi(\mathbf{r},\mathbf{r}',t-t') &= \chi_s(\mathbf{r},\mathbf{r}',t-t')+\iint d\tau\,d\tau'\iint d^3x\,d^3x'\,\chi_s(\mathbf{r},\mathbf{x},t-\tau) \\ &\quad\times\left[\frac{\delta V_H(\mathbf{x},\tau)}{\delta n(\mathbf{x}',\tau')}\frac{\delta n(\mathbf{x}',\tau')}{\delta V_0(\mathbf{r}',t')}+\frac{\delta V_{xc}(\mathbf{x},\tau)}{\delta n(\mathbf{x}',\tau')}\frac{\delta n(\mathbf{x}',\tau')}{\delta V_0(\mathbf{r}',t')}\right], \quad (31)\end{aligned}$$

which ultimately leads to the Dyson equation for the response function

$$\begin{aligned}&\chi(\mathbf{r},\mathbf{r}',t-t') = \chi_s(\mathbf{r},\mathbf{r}',t-t') \quad (32)\\ &+\iint d\tau\,d\tau'\iint d^3x\,d^3x'\,\chi_s(\mathbf{r},\mathbf{x},t-\tau)\left[\frac{\delta(\tau-\tau')}{|\mathbf{x}-\mathbf{x}'|}+f_{xc}(\mathbf{x},\mathbf{x}',\tau-\tau')\right]\chi(\mathbf{x}',\mathbf{r}',\tau'-t'),\end{aligned}$$

where we used the definition of the time-dependent Hartree potential, Eq. (20) and defined the exchange-correlation kernel

$$f_{xc}(\mathbf{x},\mathbf{x}',\tau-\tau') = \left.\frac{\delta V_{xc}(\mathbf{x},\tau)}{\delta n(\mathbf{x}',\tau')}\right|_{n_0}. \tag{33}$$

A Fourier transformation of Eq. (32) with respect to $t-t'$ transforms the equation into the corresponding Dyson equation for the response function in frequency space

$$\begin{aligned}\chi(\mathbf{r},\mathbf{r}',\omega) &= \chi_s(\mathbf{r},\mathbf{r}',\omega) \\ &\quad + \iint d^3x\, d^3x'\, \chi_s(\mathbf{r},\mathbf{x},\omega)\left[\frac{1}{|\mathbf{x}-\mathbf{x}'|} + f_{xc}(\mathbf{x},\mathbf{x}',\omega)\right]\chi(\mathbf{x}',\mathbf{r}',\omega). \end{aligned}\tag{34}$$

The response functions χ and χ_s have poles at the excitation energies of the interacting and the KS systems, respectively. From Eq. (34) we can see that the KS excitation energies, which appear in the first term on the right-hand side, are corrected by a term that depends on the Coulomb potential and the exchange-correlation kernel in order to shift them to the excitation energies of the interacting system. Therefore, calculating excitation energies by simply taking differences of the KS energy eigenvalues corresponds to setting the second term in Eq. (34) to zero.

Equation (34) was derived for the exact KS system, in practice we need to employ an approximation to the exchange-correlation potential in order to calculate χ_s and f_{xc}. Hence, linear response improves upon the excitation energies obtained from static DFT but, depending on the approximation that one uses, some errors remain. Especially the commonly used adiabatic approximations, Eq. (21), yield an exchange-correlation kernel that is independent of frequency. As a consequence, double excitations cannot be described in linear response by an adiabatic approximation [14]. We discuss this in more detail in section 3.4.

3.3 Calculating excitation energies by propagating in time

Instead of extracting spectral information from the frequency-dependent response function via the Dyson equation one can obtain the excitation energies from propagating the system in time and calculating the induced dipole moment. For a small perturbation the change in the dipole moment is given by

$$\delta\mathbf{d}(t) = \int d^3r\, \mathbf{r}\, \delta n(\mathbf{r},t) = \int dt' \iint d^3r\, d^3r'\, \mathbf{r}\chi(\mathbf{r},\mathbf{r}',t-t')\delta V(\mathbf{r}',t'). \tag{35}$$

For the calculation of optical spectra, i.e. the response of the system to an external laser, the perturbation is given as an electric field in dipole approximation and is of the form

$$\delta V(\mathbf{r},t) = \mathbf{r}\,\mathbf{E}(t) \tag{36}$$

with $\mathbf{E}(t)$ being the electric field of the laser. Introducing the dynamical polarization tensor

$$\alpha_{ij}(t-t') = -\iint d^3r d^3r'\, r_i r_j\, \chi(\mathbf{r},\mathbf{r}',t-t') \tag{37}$$

we can write the change in a component of the dipole moment for a perturbation of the form (36) as

$$\delta d_i(t) = -\sum_{j=1}^{3}\int dt' \alpha_{ij}(t-t')E_j(t'). \tag{38}$$

As for the response function we can Fourier transform this equation to frequency space and obtain

$$\delta d_i(\omega) = -\sum_{j=1}^{3} \alpha_{ij}(\omega) E_j(\omega). \tag{39}$$

The optical absorption cross section is proportional to the imaginary part of the dynamical polarizability α.
From the definition of α_{ij} we see that it has the same dependence on time or frequency as the response function itself. In other words, $\alpha_{ij}(\omega)$ has the same poles as the response function with the exception of those excitations with zero oscillator strength in the dipole approximation because the integrals in Eq. (37) then yield zero. Hence, if a transition is dipole forbidden due to some selection rules it will not appear as a pole in α_{ij}. Using Eq. (39) we can obtain the dynamical polarization tensor by running three separate calculations, one with an electric field in each of the three spatial directions, and computing the change in the dipole moment using Eq. (35). Hence, we just have to keep track of the change in the density in each of the three calculations. Equation (39) gives the impression that we have to run a separate calculation for each frequency. However, as we discuss in section 4, by choosing a specific form of the time dependence of the electric field $\mathbf{E}(t)$ we can treat all frequencies simultaneously in one calculation.

3.4 Calculating excitation energies from the Casida equation

The Casida equation [15] is an alternative way to formulate linear response as an eigenvalue problem of a frequency dependent operator. Since it is often used to calculate the excitation energies we briefly review the derivation of the equation as it is implemented in many codes.
We start from the change in the density $\delta n(\mathbf{r}, t)$ given by

$$\delta n(\mathbf{r}, \omega) = \int d^3r' \chi_s(\mathbf{r}, \mathbf{r}', \omega) \delta V_s(\mathbf{r}', \omega) \tag{40}$$

with

$$\delta V_s(\mathbf{r}', \omega) = \delta V_0(\mathbf{r}', \omega) + \int d^3x \left[\frac{1}{|\mathbf{r}' - \mathbf{x}|} + f_{xc}(\mathbf{r}, \mathbf{x}, \omega) \right] \delta n(\mathbf{x}, \omega). \tag{41}$$

Writing δV_s in this form contains the same information as the Dyson equation (34). This equation remains valid even if $\delta V_0(\mathbf{r}', \omega) = 0$. We then obtain for the change of the density

$$\delta n(\mathbf{r}, \omega) = \int d^3r' \chi_s(\mathbf{r}, \mathbf{r}', \omega) \int d^3x \left[\frac{1}{|\mathbf{r}' - \mathbf{x}|} + f_{xc}(\mathbf{r}, \mathbf{x}, \omega) \right] \delta n(\mathbf{x}, \omega). \tag{42}$$

Eq. (42) can be regarded as an eigenvalue equation for the eigenvalue 1 and eigenfunction $\delta n(\mathbf{r}, \omega)$ with a frequency dependent integral operator. Hence, the excitation frequencies that we are looking for are those frequencies for which the integral operator has an eigenvalue equal to 1. In order to solve this eigenvalue equation we first apply some algebraic manipulations to simplify the equation further.
Multiplying Eq. (42) with

$$f_{Hxc}(\mathbf{x}, \mathbf{r}, \omega) = \frac{1}{|\mathbf{x} - \mathbf{r}|} + f_{xc}(\mathbf{x}, \mathbf{r}, \omega) \tag{43}$$

and integrating over d^3r we arrive at

$$g(\mathbf{x},\omega) = \int d^3r f_{Hxc}(\mathbf{x},\mathbf{r},\omega) \int d^3r' \chi_s(\mathbf{r},\mathbf{r}',\omega) g(\mathbf{r}',\omega), \tag{44}$$

where we introduced the abbreviation

$$g(\mathbf{r},\omega) = \int d^3r' f_{Hxc}(\mathbf{r},\mathbf{r}',\omega)\delta n(\mathbf{r}',\omega). \tag{45}$$

We now expand $g(\mathbf{r},\omega)$ in the basis of product states of the static KS orbitals

$$g_{jk}(\omega) = \int d^3r\ \varphi_j^*(\mathbf{r})\varphi_k(\mathbf{r}) g(\mathbf{r},\omega). \tag{46}$$

Using the form of the KS response function, Eq. (29), we can rewrite Eq. (44) as an equation for the elements g_{jk} as

$$g_{jk}(\omega) = \sum_{j',k'} \frac{f_{k'} - f_{j'}}{\omega - (\epsilon_{j'} - \epsilon_{k'}) + i\eta} K_{jk,j'k'}(\omega) g_{j'k'}(\omega), \tag{47}$$

where the matrix elements $K_{jk,j'k'}$ are given by

$$K_{jk,j'k'}(\omega) = \iint d^3x\, d^3r\ \varphi_j^*(\mathbf{x})\varphi_k(\mathbf{x}) f_{Hxc}(\mathbf{x},\mathbf{r},\omega)\varphi_{j'}(\mathbf{r})\varphi_{k'}^*(\mathbf{r}). \tag{48}$$

Eq. (47) is only meaningful if $f_{k'} - f_{j'} \neq 0$, i.e. if we consider transitions between an occupied and an unoccupied KS orbital. In the following we drop the infinitesimal η since the true excitations ω will generally not coincide with the Kohn-Sham energy differences $\epsilon_{j'} - \epsilon_{k'}$. Hence, we are solving the eigenvalue equation (47) for frequencies ω where the denominator is non-zero even without adding the infinitesimal. We can further simplify Eq. (47) by introducing

$$\beta_{jk}(\omega) = \frac{g_{jk}(\omega)}{\omega - (\epsilon_j - \epsilon_k)} \tag{49}$$

which yields

$$\omega\beta_{jk}(\omega) = \sum_{j'k'} \left[\delta_{jj'}\delta_{kk'}(\epsilon_{j'} - \epsilon_{k'}) + (f_{k'} - f_{j'}) K_{jk,j'k'}(\omega)\right] \beta_{j'k'}(\omega). \tag{50}$$

This equation contains two different contributions: excitations where state j' is occupied and state k' is empty and those where j' is empty and k' is occupied. In order to separate those contributions we label all occupied states with j and j' and all empty states with a and a' and introduce

$$X_{ja}(\omega) = -\beta_{ja}(\omega), \quad Y_{ja}(\omega) = \beta_{aj}(\omega). \tag{51}$$

If the static KS orbitals are real, in other words if the static KS Hamiltonian has time-reversal symmetry, we can write Eq. (50) in the very compact form which is commonly referred to as the Casida equation [15]

$$\begin{pmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{B} & \mathbb{A} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \omega \begin{pmatrix} -\mathbb{1} & 0 \\ 0 & \mathbb{1} \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{52}$$

with

$$A_{ja j'a'}(\omega) = \delta_{jj'}\delta_{aa'}(\epsilon_{a'} - \epsilon_{j'}) + K_{ja,j'a'}(\omega), \quad (53)$$
$$B_{ja j'a'}(\omega) = K_{ja,j'a'}(\omega). \quad (54)$$

For a frequency independent kernel f_{xc} the matrix elements $K_{ja,j'a'}$ become real which allows for a further simplification into

$$\sum_{j'a'} \left[\delta_{jj'}\delta_{aa'}(\epsilon_{a'} - \epsilon_{j'})^2 + 2\sqrt{(\epsilon_a - \epsilon_j)(\epsilon_{a'} - \epsilon_{j'})}K_{ja,j'a'} \right] Z_{j'a'}(\omega) = \omega^2 Z_{ja}(\omega) \quad (55)$$

with

$$Z_{ja}(\omega) = \sqrt{\epsilon_a - \epsilon_j}\,(X_{ja} - Y_{ja})\,. \quad (56)$$

This alternative form significantly reduces the size of the matrix compared to the original one in Eq. (52). Therefore, this equation is implemented in computer codes rather than Eq. (52) for all adiabatic approximations.

Equation (55) reveals one important piece of information: an adiabatic approximation for $V_{xc}(\mathbf{r}, t)$ corrects the KS energy differences but does not introduce any additional excitations. For adiabatic approximations, the operator on the left-hand-side of Eq. (55) is independent of ω which turns the equation into a normal matrix eigenvalue problem. The dimension of the matrix is the number M of pairs of KS orbitals with one occupied and one unoccupied orbital. Hence, we obtain exactly M eigenvalues ω^2 leading to $2M$ eigenvalues $\pm\omega$ corresponding to the absorption and emission of a photon with frequency ω. At the same time, the number of KS energy differences $(\epsilon_{a'} - \epsilon_{j'})$ that enter the equation is M as well. Therefore, if we are using an adiabatic approximation in linear response we are only shifting the position of the excitation energies but we are not introducing additional excitations. This has dramatic consequences for excitations with double excitation character. Loosely speaking these excitations involve the change of two orbitals when going from the ground state to the excited state, although the concept of an orbital is strictly only defined for non-interacting electrons. The matrix element of the perturbation (36) between two Slater determinants, representing the KS ground and excited states, is zero if those determinants differ by more than one orbital. Hence, the KS response function does not contain any double excitations. This is also reflected in the fact that the KS energy differences $\epsilon_{j'} - \epsilon_{k'}$ are differences between two orbital energies only, there are no contributions of the form $\epsilon_j + \epsilon_{j'} - \epsilon_k - \epsilon_{k'}$ corresponding to a change of orbitals j and j' to k and k'. As a result, double excitations can only be included in linear response if the exchange-correlation kernel is frequency dependent but they do appear in the higher order KS response functions.

3.5 Comparison between time propagation and Casida equation

There is no clear advantage to calculating the excitation energies from the time propagation or the Casida equation. The time propagation includes a Fourier transformation from time to frequency space to obtain Eq. (39). In principle this transformation is done over an infinite time interval which in practice is replaced by a finite interval. In order to obtain correct results this interval needs to be chosen sufficiently large. Hence, one needs rather long propagation times in order to obtain converged results when using this method. In comparison solving the Casida equation does not need any time propagation, all quantities can directly be calculated

from the results of a static DFT calculation. However, the equation includes the summation over infinitely many unoccupied KS states, see Eq. (55). In practice this summation is truncated at some large but finite number of unoccupied states. The time propagation, however, does not need the inclusion of any unoccupied states in the calculation. Therefore, one trades the computational cost of a long time propagation for the cost of the inclusion of the unoccupied states. Depending on the specifics of the treated system one or the other method can be advantageous.

4 Applications

One of the most common applications of time-dependent DFT is the calculation of optical spectra, i.e. the information of which frequencies are absorbed by a system if it is exposed to light usually in the form of a laser. The system is originally in its ground state and the perturbing laser field is described in dipole approximation as

$$\delta V_j(\mathbf{r}, t) = E_0 r_j \delta(t - t_0), \tag{57}$$

where we run three separate calculations for the three spatial directions, i.e. $j = 1, 2, 3$. Using this type of perturbation is equivalent to multiplying each KS orbital with the same space dependent phase factor at t_0 which gives the system an initial current without changing the initial density. The perturbation excites all possible frequencies with equal intensity due to the Fourier transform to frequency space of the δ-function being a constant, i.e. $E_j(\omega) = E_0$ in Eq. (39) for all three directions. The dynamical polarization for the perturbation (57) therefore reads as

$$\alpha_{jk}(\omega) = -\frac{1}{E_0} \int_{t_0}^{\infty} dt\, \delta d_j(t) e^{i\omega t}. \tag{58}$$

While the index k does not explicitly appear on the right-hand side of this equation the induced dipole moment δd_j depends on the polarization of the applied laser field, hence, the need for three separate calculations. Of course, in a practical calculation the Fourier transformation of the dipole moment needs to be done over a finite time interval.

In the following we discuss the application of linear response TDDFT in two examples. The first is a one-dimensional 2 electron model system which has the advantage that one can solve the exact time-dependent Schrödinger equation. Hence, we can obtain the exact excitation energies for comparison. In the second application we calculate the optical spectrum of the benzene molecule. We have chosen these simple systems for pedagogical reasons only, they do not represent the capabilities of modern TDDFT calculations. In fact, TDDFT has been applied to systems as large as chromophores [16].

All the results presented in this section were obtained using the `octopus` computer code [17], a real-space pseudopotential code which can be used for both ground-state and time-dependent DFT calculations.

4.1 One-dimensional model system

In order to show the quality of a linear response TDDFT calculation we compare the excitation energies from an adiabatic LDA approximation with the exact excitations. In order to have access to the latter, we choose a one-dimensional two electron model system described by the

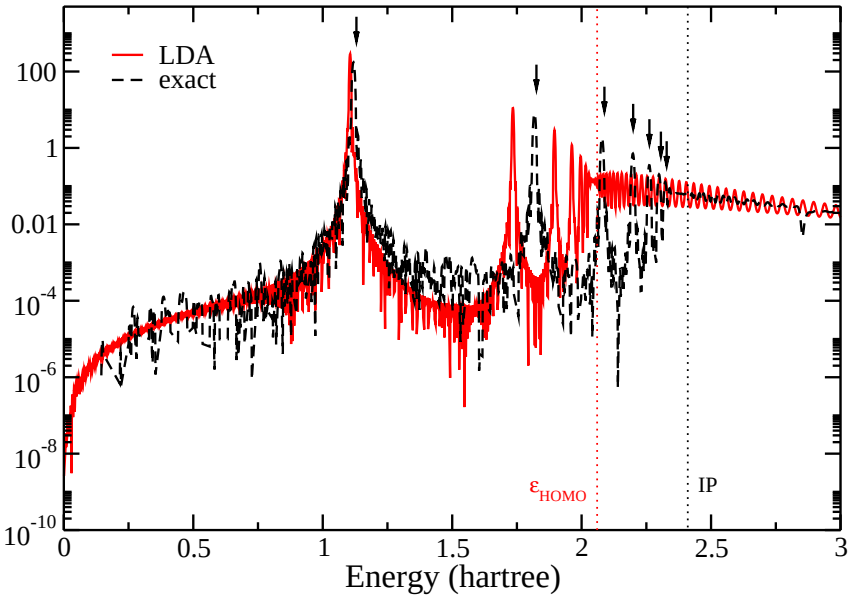


Fig. 2: *Comparison of the linear response spectrum from LDA with the exact spectrum for a one-dimensional Be^{2+} ion. Both spectra were calculated using a time propagation scheme.*

Hamiltonian

$$\hat{H} = \sum_{j=1}^{2} \left[-\frac{d^2}{2dx_j^2} - \frac{4}{\sqrt{x_j^2+1}} \right] + \frac{1}{\sqrt{(x_1-x_2)^2+1}}. \tag{59}$$

One can think about this model system as a one-dimensional Be^{2+} ion. For the Coulomb interaction, the last term in Eq. (59), we use the soft Coulomb interaction since the usual $1/|x_1 - x_2|$ introduces a singularity that is non-integrable in one dimension. For large distances between the electrons the two interactions are identical but for $x_1 = x_2$ the soft Coulomb interaction remains finite. For a detailed discussion of how to solve model systems with reduced dimensionality exactly see Ref. [18].

As we discussed in section 2.2, before running a time propagation we need to calculate the ground state of our system in order to find the initial Kohn-Sham state. We use a one-dimensional local density approximation for these calculations [19]. We then proceed using the same functional as an adiabatic approximation for the time-dependent exchange-correlation potential. For the calculation of the spectrum we apply a δ-type perturbation, Eq. (57), with $E_0 = 10^{-4}$ Ha/bohr and propagate for a total propagation time of 10^3 a.u. The resulting spectrum, is shown in Fig. 2. As a guide to the eye we have also included a line for the energy ϵ_{HOMO} of the highest occupied Kohn-Sham level and the IP, the exact ionization potential. They both mark the onset of the continuum part of the spectra. In the discrete part of the spectrum, we can clearly identify 7 peaks in the exact spectrum from which only 5 are visible in the LDA spectrum. This is a well-known failure of the LDA approximation, it does not yield the Rydberg series in the spectrum due to the exchange-correlation potential decaying exponentially in LDA rather than with the correct $-1/r$ behavior for $r \to \infty$. This failure is also an example where the approximation to the static exchange-correlation potential influences the quality of the spectrum. The adiabatic exact exchange functional, for example, does yield a Rydberg series and produces all 7 peaks [19]. The almost equally spaced excitations in the LDA spectrum

	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7
LDA	1.10	1.74	1.90	1.96	2.00	-	-
exact	1.12	1.81	2.08	2.19	2.26	2.29	2.32
$\Delta\epsilon_{ij}$	1.02	1.73	1.89	1.96	2.00	-	-

Table 1: *Excitation energies from linear response of the 1D Be^{2+} atom corresponding to the spectra in Fig. 2. For the LDA excitations we also list the corresponding Kohn-Sham energy differences $\Delta\epsilon_{ij} = \epsilon_i - \epsilon_j$. All numbers are given in Hartree.*

above 2.1 Ha represent excitations to the continuum. Since we are performing our calculations in a large but still finite box our continuum states are represented as box states, i.e. with zero boundary conditions at the box edge. In the time-dependent calculation we need to include absorbing boundary conditions in order to avoid the reflection of parts of the density that reach the boundary during the propagation. The continuum part of the spectrum is influenced by the details of those absorbing boundary conditions. The small dip at 2.8 Ha that is visible in the exact spectrum but missing from the LDA one corresponds to a Fano resonance, i.e. a decay of a localized excited state to a continuum state. Since the localized excited state in this case is a double excitation it cannot be reproduced by any adiabatic functional as we discussed in section 3.4.

The excitation energies extracted from Fig. 2 are given in table 1. For the LDA excitations we also list the corresponding differences of the Kohn-Sham energy eigenvalues, i.e. the excitations that are given by χ_s. As we can see, the corrections from f_{Hxc} in the Dyson equation (34) are positive for all excitations and larger for the lower lying excitations than for the higher ones.

4.2 Benzene

For the calculation of the benzene molecule we employ the LDA functional of Perdew and Zunger [20] modified to improve the matching between the high and low density regime. We then use the same functional in an adiabatic fashion for the time-dependent calculation. We run three separate calculations with the polarization of the electric field in the δ-kick pointing in one of the three spatial directions for each calculation. From the resulting change in the dipole moment, we construct the whole tensor for the dynamical polarizability. The trace of the imaginary part of this tensor yields the average cross section for the absorption of a photon, i.e. it yields the excitation energies of the system. We plot this cross section in Fig. 4.2. The excitation at 6.9 eV corresponds to a $\pi - \pi^*$ transition and agrees perfectly with experiment [21]. All other excitations that are visible in the figure lie in the continuum. They are not well represented because we run the calculation in a finite box with absorbing boundary conditions. These are, as in the one-dimensional example before, necessary to avoid a reflection of those parts of the density that reach the edge of the box in the time-dependent calculation. The continuum part of the spectrum is modified by the details of the absorbing boundary conditions, hence, it is not representative for the benzene molecule.

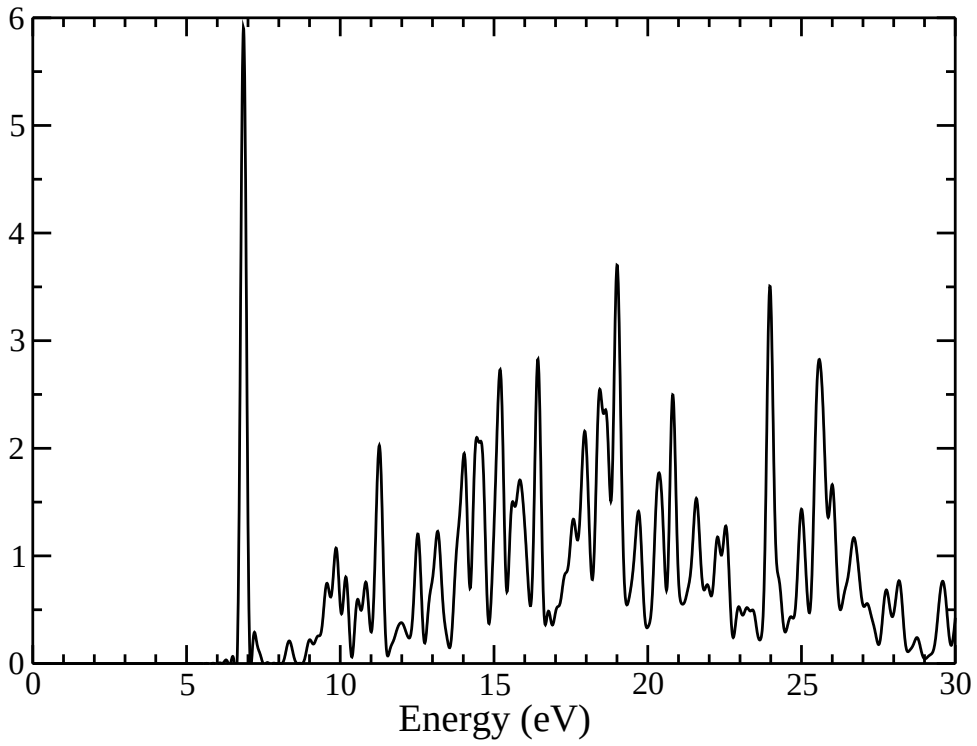


Fig. 3: *Spectrum of a benzene molecule obtained from a time propagation. The excitations above 8 eV lie in the continuum, the peak at 6.9 eV agrees very well with experiment [21].*

References

[1] C.A. Ullrich, *Time-Dependent Density-Functional Theory, Concepts and Applications*, Oxford University Press (2012).

[2] M.A.L. Marques, N.T. Maitra, F.M.S. Nogueira, E.K.U. Gross, A. Rubio (Eds.), *Fundamentals of Time-Dependent Density Functional Theory*, Springer-Verlag Berlin Heidelberg (2012).

[3] E. Runge, E.K.U. Gross, Phys. Rev. Lett. **52**, 997 (1984).

[4] M. Ruggenthaler, R. van Leeuwen, Europhys. Lett. **95**, 13001 (2011).

[5] N.T. Maitra, I. Souza, K. Burke, Phys. Rev. B **68**, 045109 (2003).

[6] G. Vignale, Phys. Rev. B **70**, 201102 (2004).

[7] R. van Leeuwen, Phys. Rev. Lett. **82**, 3863 (1999).

[8] J.E. Harriman, Phys. Rev. A **24**, 680 (1981).

[9] P. Hohenberg, W. Kohn, Phys. Rev. **136**, B864 (1964).

[10] N.T. Maitra, K. Burke, C. Woodward, Phys. Rev. Lett. **89**, 023002 (2002).

[11] N.T. Maitra, K. Burke, Phys. Rev. A **63**, 042501 (2001), N.T. Maitra, K. Burke, Phys. Rev. A **64**, 039901 (2001).

[12] A. Castro, M.A.L. Marques, A. Rubio, J. Chem. Phys. **121**, 3425 (2004).

[13] A.L. Fetter, J.D. Walecka, *Quantum Theory of Many-Particle Systems,* (Dover Publications, 2003).

[14] N.T. Maitra, F. Zhang, R.J. Cave, K. Burke, J. Chem. Phys. **120**, 5932 (2004).

[15] M.E. Casida, *Time-dependent density functional response theory for molecules* in: D.E. Chong (Ed.) *Recent Advances in Density Functional Methods. Recent Advances in Computational Chemistry*, vol. 1, pp. 155-192, (World Scientific, Singapore 1995).

[16] M.A.L. Marques, A. Castro, G.F. Bertsch, A. Rubio, Comput. Phys. Comm. **151**, 60 (2003); A. Castro, M.A.L. Marques, J.A. Alonso, A. Rubio, J. of Comput. Theor. Nanoscience **1**, 231 (2004).

[17] A. Castro, et al., Phys. Status Solidi B, **243**, 2465 (2006); X. Andrade, et al., J. Phys.: Cond. Matt. **24**, 233202 (2012).

[18] N. Helbig et al., Chem. Phys. **391**, 1 (2011).

[19] N. Helbig et al., Phys. Rev. A **83**, 032503 (2011).

[20] J.P. Perdew, A. Zunger, Phys. Rev. B **23**, 5048 (1981).

[21] K. Yabana, G.F. Bertsch, Int. J. Quant. Chem. **75**, 55 (1999).

A 8 DFT+*U* method and *ab-initio* determination of the Coulomb interaction parameter *U* [1]

Ersoy Şaşıoğlu
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

One of the main goals of condensed matter physics is the solution of the many-body Schrödinger equation, $H\psi = \epsilon\psi$, with the Hamiltonian

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,\alpha}\frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} - \sum_\alpha \frac{1}{2M_\alpha}\nabla_\alpha^2 + \frac{1}{2}\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}, \quad (1)$$

where $\mathbf{r}_i$ ($\mathbf{R}_\alpha$) are the coordinates of the N_e electrons (N_n nuclei), Z_α is the atomic number of the nuclei, and M_α the corresponding nuclear masses. The Born-Oppenheimer approximation allows a decoupling of the motion of the electrons from that of the nuclei resulting in an electronic Hamiltonian of the form

$$\begin{aligned}\hat{H}_e &= -\frac{1}{2}\sum_i \nabla_i^2 + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,\alpha}\frac{Z_\alpha}{|\mathbf{r}_i - \mathbf{R}_\alpha|} \\ &= \hat{T}_e + \hat{V}_{ee} + \hat{V}_{en}\,, \end{aligned} \quad (2)$$

where $\hat{T}_e$ is the kinetic energy operator, $\hat{V}_{ee}$ is the electron-electron interaction operator, and the last term, i.e., the periodic attractive lattice potential $\hat{V}_{en}$ is often denoted as $\hat{V}_{ext}$ to emphasize the pure electronic character. The remaining lattice part, which will not be covered here, describes the motion of the nuclei and gives rise to phonons.

The Schrödinger equation for the electrons, $H_e\psi = \epsilon\psi$, can be solved exactly in the non-interacting limit, i.e., for $\hat{V}_{ee} = 0$. On the other hand, the electron-electron interaction $\hat{V}_{ee}$ makes the problem extremely complicated. Except for small systems, e.g., for atoms and molecules, various simplifications and approximations are needed in order to treat the problem. In solid state theory there are two main approaches: i) Many-body model Hamiltonians such as Hubbard model, Anderson *s-d* model etc., which are all based on a set of adjustable parameters and take into account a restricted number of electronic states in a small energy window around the Fermi level (see Fig. 1). ii) *Ab-initio* electronic structure methods for which the only input parameters are the atomic numbers, positions of atoms, and the number of electrons in the system. The most widely used *ab-initio* method is based on density functional theory (DFT) [1, 2], which maps the many-electron problem onto a one-electron one.

The scope of this lecture is to combine DFT and many-body methods like the Hubbard model. In Section 2, we will briefly discuss DFT and its limits. Section 3 focuses on combining DFT and many-body model Hamiltonians, in particular the DFT+*U* method will be presented. As a very efficient scheme to estimate Coulomb interaction parameter U we introduce in Section 4 the constrained random-phase approximation (cRPA). Applications of cRPA to various materials will be presented in Section 5.

2 DFT in a nutshell

Density functional theory is described in detail in the lecture A2 by M. Ležaić. Here, we will briefly overview DFT to make the present chapter self-contained. DFT is based on the Hohenberg-Kohn theorem [3], which states that there is a one-to-one correspondence between the ground-state electron density $n(\mathbf{r})$ of the inhomogeneous interacting electron system and the external potential $v_{\text{ext}}(\mathbf{r})$ acting on it. Thus, for any material the ground state total energy

can be regarded as a functional of the electron density, $E[n]$, which can be minimized by the ground-state density. The total energy $E[n]$ can be partitioned into

$$E[n] = F[n] + \int d\mathbf{r} n(\mathbf{r})\, v_{\text{ext}}(\mathbf{r})\,, \tag{3}$$

where $F[n] = T_e[n] + E_{ee}[n]$ is the sum of the kinetic and electron-electron interaction energy. The latter is an unknown universal functional. The difficulty lies in the fact that the determination of $n(\mathbf{r})$ requires, in principle, the solution of the many-body problem. However, Kohn and Sham [4] have shown that $n(\mathbf{r})$ can be obtained from the solution of the Schrödinger equation of a fictitious non-interacting electron system moving in an effective potential $v_{\text{eff}}(\mathbf{r})$. The effective potential $v_{\text{eff}}(\mathbf{r})$ is chosen in such a way that the ground-state density $n_0(\mathbf{r})$ equals the ground-state density, $n(\mathbf{r})$ of the interacting system

$$n(\mathbf{r}) = n_0(\mathbf{r}) = \sum_n^{\text{occ}} |\psi_n(\mathbf{r})|^2\,. \tag{4}$$

The universal functional $F[n]$ can be now written as

$$\begin{aligned} F[n] &= T_0[n] + E_{\text{H}}[n] + E_{\text{xc}}[n] \\ &= T_0[n] + \iint d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + E_{\text{xc}}[n]\,, \end{aligned} \tag{5}$$

where $T_0[n]$ is the kinetic energy of the non-interacting auxiliary system, E_{H} the classical electrostatic Hartree energy, and $E_{\text{xc}}[n]$ is the so-called exchange-correlation energy,

$$E_{\text{xc}} = E_{\text{ee}}[n] - E_{\text{H}}[n] + T_{\text{e}}[n] - T_0[n]\,. \tag{6}$$

The exchange-correlation energy $E_{\text{xc}}[n]$ comprise in principle all complicated many electron effects. It has to describe the reduction of the Coulomb interaction between the electrons in the same spin state due to the Pauli exclusion principle (so-called exchange interaction) and correlation effects stemming from the correlated movement of the electrons.
The Kohn-Sham equation can be obtained by minimizing the total energy with respect to ψ_n using the constraint $\langle\psi_n|\psi_{n'}\rangle = \delta_{n,n'}$

$$\begin{aligned} H_e^0(\mathbf{r})\psi_n(\mathbf{r}) &= \big[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r})\big]\psi_n(\mathbf{r}) \\ &= \epsilon_n\psi_n(\mathbf{r})\,, \end{aligned} \tag{7}$$

where the eigenvalues ϵ_n are strictly speaking the Lagrange multipliers, which guarantees the orthogonality of the orbitals and the effective potential $v_{\text{eff}}(\mathbf{r})$ is given by

$$\begin{aligned} v_{\text{eff}}(r) &= -\sum_\alpha \frac{Z_\alpha}{|\mathbf{r}-\mathbf{R}_\alpha|} + \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta E_{xc}[n]}{\delta n} \\ &= v_{\text{ext}}(\mathbf{r}) + v_{\text{H}}(\mathbf{r}) + v_{xc}(\mathbf{r})\,. \end{aligned} \tag{8}$$

The ground-state electron density $n(\mathbf{r})$ can be obtained by solving Eq. (7) self-consistently. In conclusion, in DFT the interacting many-electron problem is reduced to a non-interacting single electron problem. The non-interacting electrons move in an effective potential $v_{\text{eff}}(\mathbf{r})$ (mean-field) created by the nuclei $v_{\text{ext}}(\mathbf{r})$, electro-static Hartree potential $v_{\text{H}}(\mathbf{r})$ created by the charge

of the electrons, and the exchange-correlation potential $v_{\rm xc}(\mathbf{r})$. DFT would be an exact theory if $E_{\rm xc}[n]$ was known. However, the exchange-correlation functional $E_{\rm xc}[n]$ is an unknown. In practice, it is therefore necessary to use approximate forms. The most popular exchange-correlation functional is the local-density approximation (LDA) [5] and its extensions such as the generalized gradient approximation (GGA), meta-GGA, etc [6]. In LDA the $E_{\rm xc}[n]$ depends locally on the value of the electronic density at each point in space

$$E_{\rm xc}[n] = \int d\mathbf{r}\, n(\mathbf{r}) \varepsilon_{\rm xc}^{\rm LDA}(n(\mathbf{r}))\,, \tag{9}$$

where $\varepsilon_{\rm xc}^{\rm LDA}(n(\mathbf{r}))$ is the contribution of the exchange and correlation effects to the total energy of the homogenous interacting electron gas with a constant density that is identical to the local density $n(\mathbf{r})$ of the inhomogeneous system. For spin-polarized systems one uses the local spin density approximation (LSDA) [7]

$$E_{\rm xc}[n^{\uparrow}, n^{\downarrow}] = \int d\mathbf{r}\, n(\mathbf{r}) \varepsilon_{\rm xc}^{\rm LDA}(n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r}))\,, \tag{10}$$

where $\varepsilon_{\rm xc}^{\rm LDA}(n^{\uparrow}(\mathbf{r}), n^{\downarrow}(\mathbf{r}))$ is the spin-dependent exchange correlation energy in a spin-polarized electron gas having densities $n^{\uparrow}(\mathbf{r})$ and $n^{\downarrow}(\mathbf{r})$ for spin-up and spin-down projections, respectively. Then the exchange-correlation potential becomes spin dependent,

$$v_{\rm xc}^{\sigma}(\mathbf{r}) = \frac{\delta E_{\rm xc}[n^{\uparrow}, n^{\downarrow}]}{\delta n^{\sigma}(\mathbf{r})}\,. \tag{11}$$

The L(S)DA is justified for systems with a slowly varying spatial electron density $n(\mathbf{r})$. Many successes of modern condensed-matter physics are related to the development of the DFT-L(S)DA. Through the years DFT have provided insight not only in condensed-matter physics but also in chemistry, materials science, and even in biology. Because of this the DFT has become the standard model for electronic structure calculations [1, 2]. Despite the great success of DFT in the description of the electronic and magnetic structure of weakly correlated materials, it however fails for strongly correlated materials as will be discussed in the following sections.

2.1 Strength of electronic correlations in solids

The electrons lying energetically close to the Fermi level play a decisive role in determining the properties of a solid such as the crystal structure, magnetism, conductivity, optical absorbtion, etc. As a simple example we present in Fig. 1 the electronic band structure of $SrVO_3$. As seen the states around the Fermi level correspond to the vanadium 3*d* orbitals with t_{2g} character, which are isolated from the rest of the bands and are mainly responsible for electronic properties of the material. The correlation strength for the electrons around the Fermi level appears to be an important parameter in classification of the materials and thus the choice of the methods to study their electronic structure. Correlation strength in solids is defined as the ratio of the average Coulomb interaction (Hubbard U parameter) between localized electrons and the band width W (or average kinetic energy), i.e., U/W. Solids can be divided into three categories depending on their U/W ratio.

- $\frac{U}{W} < 1$: Weak correlations
- $\frac{U}{W} \sim 1$: Intermediate correlations

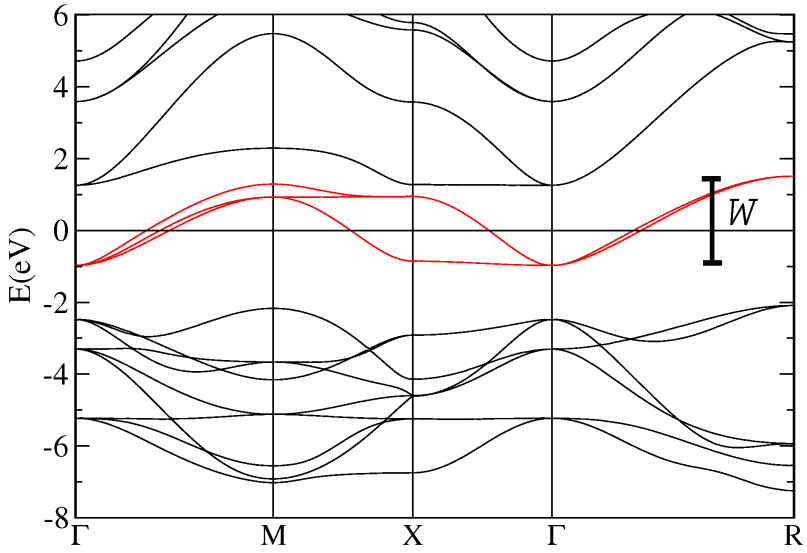


Fig. 1: *Band structure of the $SrVO_3$. The red bands correspond to the vanadium 3d orbitals with t_{2g} character, which are isolated from the rest of the bands and are mainly responsible for electronic properties of the material. The W denotes the band width of correlated states around the Fermi energy*

- $\frac{U}{W} > 1$: Strong correlations

Materials with $U/W < 1$ shows free electron-like behavior, i.,e, valence electrons are itinerant while those with $U/W > 1$ posses atomic-like features, i.e., valence electrons are localized in atomic sites. Very rich physics appears in between, e.g., for systems with $U/W \sim 1$, in which kinetic energy and Coulomb potential energy compete making these systems sensitive to external perturbations. For instance, applied pressure can induce metal-insulator phase transition in such materials. Magnetic phase transition, orbital order, superconductivity, etc, are among the other examples of emergent phenomena in correlated materials. In Fig. 2 we present some selected materials ranging from from weak to strong correlations. As seen the 3*d* transition metals (TMs) and their compounds are positioned on both sides of $U/W \sim 1$ value, which means that these materials show diverse electronic and magnetic properties. The U/W values are also important in the choice of the methods to study electronic structure of the corresponding systems [8, 9, 10]. The two main approaches in solid state theory mentioned at the beginning are best appropriate for distinctive correlation strengths. While *ab-initio* electronic structure methods such as DFT-LDA is generally used for materials from weak to intermediate correlations, model Hamiltonians such as Hubbard model is much more suitable for systems from intermediate to strong correlations.

2.2 Limits of the DFT method

As mentioned above, despite the great success of DFT-L(S)DA in the description of the electronic and magnetic structure of weakly correlated materials, i.e., for systems with $U/W < 1$, the exchange-correlation functional [see Eq. (9)] based on the homogenous electron gas is not adequate for materials with narrow energy bands i.e., for systems ranging from intermediate to strong correlations (see Fig. 2). In such materials valence electrons preserve a significant de-

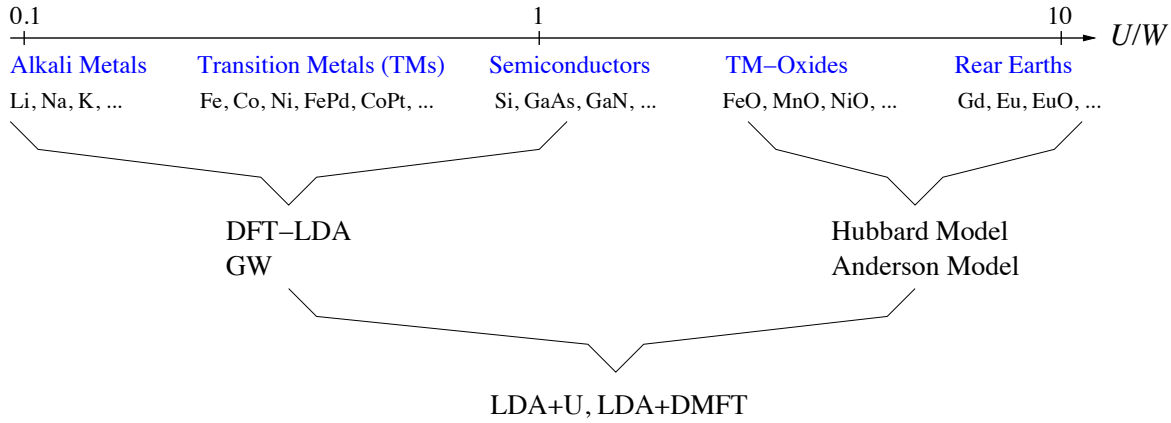


Fig. 2: *Correlation strength U/W for selected materials. Note that a logarithmic scale is used for U/W. Also shown are the different approaches in solid state theory and their range of applicability in terms of U/W value.*

gree of their atomic orbital nature and thus DFT-L(S)DA might give qualitatively wrong results. Classical examples of this failure are Mott insulators [11] such as TM-oxides (NiO, MnO, etc.) The insulating state stems from the strong Coulomb repulsion between electrons residing in partially filled *d* orbitals. When the kinetic energy gain due to the hopping of *d* (*f*) electrons from site to site is smaller than the energy loss due to the Coulomb repulsion between two electrons on the same site, i.e., for $U/W > 1$, electron localization takes place leading to an insulating state, which cannot be described properly by L(S)DA.
In DFT-L(S)DA the electron moves in an effective field created by all the other electrons in the system including itself, which gives rise to the *self-interaction* problem. This self-interaction is partially taken care of by the exchange correlation functional in L(S)DA, but not completely like in the GW approximation or the Hartree-Fock method. As a consequence of this residual self-interaction the L(S)DA underestimates the band gaps of semiconductors and insulators [12]. Another consequence of the self-interaction error is that DFT-L(S)DA is not able to treat charge ordering effects in correlated materials [13, 14, 15]. For instance, for Fe_3O_4 the DFT calculations within L(S)DA [14] give rise to a metallic state without charge ordering in contrast to the experimentally observed charge ordered insulating ground state. However, the DFT+*U* method as will be introduced below, which takes into account correlation effects, gives the correct ground state [15].

3 Combining DFT and many-body methods

For a long time, DFT and many-body model Hamiltonian approaches have been separate and complementary methods. This has drastically changed with the advent of the LDA+U and LDA+DMFT methods, where the DMFT stands for the dynamical mean-field theory [16, 17, 18]. The former approach, i.e., the LDA+*U* method [18], an early attempt to correct the LDA functional by introducing a simple mean-field-like Hubbard *U* term for localized *d* or *f* states, and today routinely applied to a broad spectrum of systems, can be regarded as the static limit of LDA+DMFT approach. Both LDA+*U* and LDA+DMFT have in common that they rely on the Coulomb interaction *U* as an input parameter. Frequently the exact value of *U* is unknown which impedes the predictive power of these approaches.

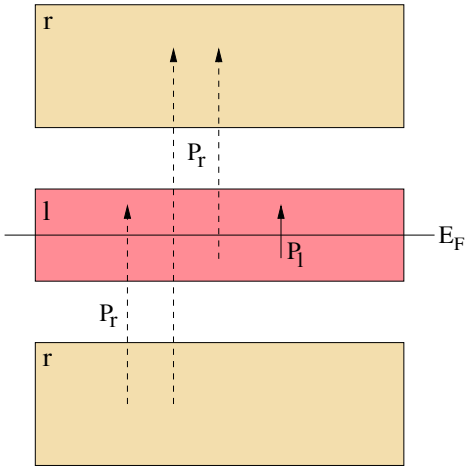


Fig. 3: *Schematic representation of the strongly correlated l-subspace and weakly correlated r-subspace. The P_l and P_r denote polarization functions for the l-subspace and the rest. Although the P_l is confined to the transitions within the l-subspace, P_r may contain transitions between the l- and r-subspaces.*

In order to combine DFT and many-body methods it is necessary to define an optimum Hamiltonian for the system under study and determine its parameters from DFT. To this end we divide the Hilbert space of the crystal electronic states into two subspaces as depicted schematically in Fig. 3: i) *weakly correlated subspace* (r-subspace) for which $U/W < 1$, i.e, L(S)DA is good approximation and ii) *strongly correlated subspace* (l-subspace) for which $U/W > 1$, i.e, L(S)DA fails completely. For the strongly correlated subspace the Coulomb interaction term of the Hamiltonian should be written in full form, while the one-particle part can be described by the L(S)DA. Another important issue is the definition of atomic-like orbitals for the electrons in the strongly correlated subspace. The suitable choice for such orbitals are the maximally localized Wannier functions (MLWFs) [19, 20], which resemble atomic orbitals, i.e., being centered on atoms and decay with increasing distance from the nuclei. Furthermore, they form an orthonormal basis set. The MLWFs $w^{\sigma}_{in}(\mathbf{r})$ with orbital index n and spin σ at site i are defined as Fourier transforms of the Bloch states $\varphi^{\sigma}_{\mathbf{k}m}(\mathbf{r})$ according to

$$w^{\sigma}_{in}(\mathbf{r}) = \frac{1}{N}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}_i} \sum_{m} T^{\sigma(\mathbf{k})}_{i,mn}\varphi^{\sigma}_{\mathbf{k}m}(\mathbf{r}) \,, \tag{12}$$

where N is the number of discrete $\mathbf{k}$ points in the full Brillouin zone and $T^{\sigma(\mathbf{k})}_{i,mn}$ denotes the transformation matrices. The latter are determined by minimizing the spread

$$\Omega = \sum_{i,n,\sigma} \left(\langle w^{\sigma}_{in}|r^2|w^{\sigma}_{in}\rangle - \langle w^{\sigma}_{in}|\mathbf{r}|w^{\sigma}_{in}\rangle^2 \right) , \tag{13}$$

where the sum runs over all Wannier functions. An efficient algorithm for minimizing the spread Ω was proposed by Marzari and Vanderbilt [19] for isolated groups of bands and later extended to entangled energy bands [20].
The many-body Hamiltonian (or generalized Hubbard model) for the strongly correlated subspace is given by

$$\hat{H}_e = \hat{H}^{\mathrm{DFT}} + \hat{U}^l - \hat{H}^l_{\mathrm{DC}} \,, \tag{14}$$

where $\hat{H}^{\text{DFT}}$ is the DFT part of the Hamiltonian, $\hat{U}^l$ the effective Coulomb interaction, and $\hat{H}_{\text{DC}}$ is the double counting term. The $\hat{H}_{\text{DC}}$ is necessary in order to avoid the double counting of the Coulomb and exchange interactions which are included in the L(S)DA energy functional. The DFT part of the Hamiltonian can be expressed as

$$\hat{H}^{\text{DFT}} = -\sum_{\sigma} \sum_{in,i'n'} t^{i,i'}_{n,n'} c^{\dagger}_{in\sigma} c_{i'n'\sigma} \,, \tag{15}$$

where $c^{\dagger}_{in\sigma}$ ($c_{in\sigma}$) creates (destroys) and electron with spin σ in orbital n at site i, and

$$t^{i,i'}_{n,n'} = -\int d\mathbf{r} w^{\sigma *}_{in}(\mathbf{r}) \big[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\mathbf{r}) \big] w^{\sigma}_{i'n'}(\mathbf{r}) \,. \tag{16}$$

The on-site terms ($i = i'$) give rise to crystal-field matrix while the off-site terms ($i \neq i'$) are the hopping integrals. The on-site Coulomb interaction ($i = i' = i'' = i'''$) $\hat{U}$ is given by

$$\hat{U}^l = \frac{1}{2} \sum_{i} \sum_{\sigma,\sigma'} \sum_{n,n',n'',n'''} U^l_{n,n'';n',n'''} c^{\dagger}_{in\sigma} c^{\dagger}_{in''\sigma'} c_{in'''\sigma'} c_{in'\sigma} \tag{17}$$

with

$$U^l_{n,n'';n',n'''} = \iint d\mathbf{r} d\mathbf{r}' w^{\sigma *}_{in}(\mathbf{r}) w^{\sigma' *}_{in''}(\mathbf{r}') U(\mathbf{r},\mathbf{r}') w^{\sigma'}_{in'''}(\mathbf{r}') w^{\sigma}_{in'}(\mathbf{r}) \,, \tag{18}$$

where $U(\mathbf{r},\mathbf{r}')$ is the partially screened Coulomb interaction, which is screened by the electrons of the weakly correlated subspace. Despite the restriction to on-site Coulomb interaction the Hamiltonian still describes the many-body problem. However, as will be seen later such a problem can be solved by DFT+U method, which treats correlations on a mean-field level, or by more sophisticated methods like LDA+DMFT approach, which will not be considered here.

3.1 Double counting problem

The double counting problem stems from the fact that two different approaches are employed for the calculation of the Coulomb interaction energy in DFT and model Hamiltonian methods. When the Coulomb interaction term is added to the one-electron DFT-L(S)DA Hamiltonian [see Eq. (14)] then it is necessary to subtract the part of the Coulomb energy that has already been taken into account by DFT-L(S)DA. Unfortunately, there is no rigorous way to take this correction into account because the Coulomb interaction energy in DFT-L(S)DA is calculated as functional of the charge density distribution, however in model Hamiltonian approaches this energy is expressed in terms of sum of pair interactions for the electrons on atomic orbitals.
To define $\hat{H}_{\text{DC}}$ we have to express the Coulomb interaction energy in DFT-L(S)DA in the formalism of many-body model Hamiltonian. In DFT-L(S)DA this energy is a functional of electron density, which is defined by total number of localized electrons N^l in strongly correlated subspace. Then it is reasonable to assume that in DFT-L(S)DA the double counting Coulomb interaction energy is simply a function of N^l. Neglecting the exchange contribution for a while the double counting energy E^l_{DC} can be written as

$$E^l_{\text{DC}} = \frac{1}{2} U_{av} N^l (N^l - 1) \,, \tag{19}$$

where $N^l = N^l_\uparrow + N^l_\downarrow = \sum_{m,\sigma} \hat{n}_{im\sigma}$ is the number of localized electrons per site and U_{av} is the average on-site Coulomb interaction parameter (or Hubbard U). Including the Coulomb exchange term J_{av} one gets the following expression for the double counting Coulomb interaction energy

$$E^l_{\rm DC} = \frac{1}{2}U_{av}N^l(N^l-1) - \frac{1}{2}J_{av}\sum_\sigma N^l_\sigma(N^l_\sigma - 1)\,. \tag{20}$$

This is known as the fully localized limit since orbitals are assumed to be fully occupied. According to Janak's theorem [21] the correction to the atomic orbital energies $\epsilon^\sigma_{\rm DC}$ is given by the derivative of the double counting energy with respect to the corresponding state occupancy $\hat{n}_{im\sigma}$

$$\epsilon^\sigma_{\rm DC} = \frac{\delta E^l_{\rm DC}}{\delta \hat{n}_{im\sigma}} = U_{av}\big(N^l - \frac{1}{2}\big) - J_{av}\big(N^l_\sigma - \frac{1}{2}\big)\,, \tag{21}$$

thus the term in the many-body Hamiltonian responsible for the double counting correction $\hat{H}^l_{\rm DC}$ can be written as

$$\hat{H}^l_{\rm DC} = \sum_{im\sigma} \epsilon^\sigma_{DC}\hat{n}_{im\sigma}\,. \tag{22}$$

The difference $\hat{U}^l - \hat{H}^l_{\rm DC}$ gives the short-range (on-site) many body correction to the DFT-L(S)DA for the strongly correlated subspace.

3.2 Coulomb interaction

Coulomb interaction plays a central role in the construction of the model Hamiltonians and the study of the strongly correlated materials. Calculation of the Coulomb interaction between the localized electrons in the strongly correlated subspace is a difficult task because this interaction is substantially screened thorough the transitions between two subspaces (see Fig. 3). Thus, here we will first consider the Coulomb interaction full Hilbert space, which is unscreened and renormalization (screening) of the Coulomb interaction in strongly correlated subspace will be discussed in Section 4. It is worth to note that the screened Coulomb interaction has the same form as bare one. It is advantageous to use atomic orbitals $\psi_{nlm}(\mathbf{r}) = R_{nl}(r)Y^l_m(\theta,\phi)$ instead of Wannier functions since in the former basis the bare Coulomb interaction can be parameterized in terms of Slater integrals [22].

In the atomic limit one can express bare Coulomb interaction U via complex spherical harmonics and Slater integral parameters F^k [22]. Using the spherical coordinates

$$\mathbf{r}_i = r_i(\sin\theta_i\cos\phi_i, \sin\theta_i\sin\phi_i, \cos\theta_i) \tag{23}$$

one can express the bare Coulomb interaction as

$$\frac{1}{|\mathbf{r}-\mathbf{r}'|} = \sum_{k=0}^{\infty}\frac{r^k_<}{r^k_>}\frac{4\pi}{2k+1}\sum_{q=-k}^{k} Y^{k*}_q(\theta,\phi)Y^k_q(\theta',\phi')\,, \tag{24}$$

where $r_< = r$, $r_> = r'$ if $r' > r$ and vice versa. Using Eq. (24) one can write the bare Coulomb matrix as

$$U_{m,m'';m',m'''} = \sum_{k=0}^{2l} a_k(m,m'';m',m''')F^k\,. \tag{25}$$

In Eq. (25) a_k are the Clebsch-Gordon coefficients

$$a_k(m,m'';m',m''') = \frac{4\pi}{2k+1}\sum_{q=-k}^{k}\langle lm|Y_q^k|lm'\rangle\langle lm''|Y_q^{k*}|lm'''\rangle\,, \tag{26}$$

where $\langle lm|Y_q^k|lm'\rangle$ can be expressed via the integrals over the product of three spherical harmonics

$$\langle lm|Y_q^k|lm'\rangle = \int d\mathbf{r}\, Y_l^m(\mathbf{r})Y_q^k(\mathbf{r})Y_l^{m'}(\mathbf{r})\,, \tag{27}$$

which could be calculated analytically [23] and F^k are the radial Slater integrals with $k = 0, 2, 4$ for *d* states and $k = 0, 2, 4, 6$ for *f* states. For a given n and l the F^k is given by

$$F^k = \int_0^\infty dr r^2 \int_0^\infty dr' r'^2 R_{nl}^2(r)\frac{r_<^k}{r_>^{k+1}}R_{nl}^2\,. \tag{28}$$

In Eq. (25) the two-index terms, i.e., $m = m'$ and $m'' = m'''$ ($m = m'''$ and $m' = m''$) are the most important Coulomb integrals. The rest of matrix elements vanishes due to symmetry reasons in the atomic limit. The two-index direct ($U_{m,m'}$) and exchange ($J_{m,m'}$) Coulomb matrix elements are given by

$$\begin{aligned} U_{m,m'} &= U^b_{m,m';m,m'} = \sum_{k=0}^{2l} a_k(m,m;m',m')F^k\,, \\ J_{m,m'} &= U^b_{m,m';m',m} = \sum_{k=0}^{2l} a_k(m,m';m',m)F^k\,. \end{aligned} \tag{29}$$

Using the properties of the Clebsch-Gordon coefficients the following relation can be found between the Slater integrals F^k and the average Coulomb interaction parameters U_{av} and J_{av} [23]

$$U_{av} = \frac{1}{(2l+1)^2}\sum_{m,m'} U_{m,m'} = F^0\,, \tag{30}$$

$$U_{av} - J_{av} = \frac{1}{2l(2l+1)}\sum_{m\neq m'}(U_{m,m'} - J_{m,m'}) \tag{31}$$

with

$$J_{av} = \frac{1}{2l(2l+1)}\sum_{m\neq m'} J_{m,m'}\,. \tag{32}$$

Then in terms of the Slater parameters the J_{av} for 3*d* and 4*f* orbitals are given by

$$\begin{aligned} J_{av} &= \frac{1}{14}(F^2+F^4) \qquad \text{for} \quad l=2\,, \\ J_{av} &= \frac{1}{6435}(286F^2+195F^4+250F^6) \quad \text{for} \quad l=3\,. \end{aligned} \tag{33}$$

Slater integrals represent the radial part of the Coulomb interaction, which is mostly affected by the screening effects. The F^k are usually calculated from the Coulomb interaction parameters U_{av} and J_{av} that are obtained either from experiments or DFT calculations. The simplest

case is the integral $F^0 = U$. Other Slater integrals with $k \neq 0$ are calculated via the exchange parameter. For hydrogen-like 3*d* orbitals $F^4/F^2 = 15/23$ [24], while for 4*f* orbitals $F^4/F^2 = 451/675$ and $F^6/F^2 = 1001/2025$ [25]. For atomic *d* orbitals the U_{av} is very large (15-25 eV), however screening effects reduce it drastically as will be seen in Section 4.
If we neglect all terms in Coulomb matrix except the direct and exchange contributions, i.e., only density-density terms remain, the Coulomb interaction term in many-body Hamiltonian [see Eq. (14)] takes a simple form

$$\hat{U}^l = \frac{1}{2}\sum_{i,\sigma}\sum_{m,m'} U_{m,m'}\hat{n}_{im\sigma}\hat{n}_{im'-\sigma} + \frac{1}{2}\sum_{i,\sigma}\sum_{m\neq m'}(U_{m,m'} - J_{m,m'})\hat{n}_{im\sigma}\hat{n}_{im'\sigma}\,, \tag{34}$$

where $\hat{n}_{im\sigma} = c^\dagger_{im\sigma}c_{im\sigma}$ is the particle number operator. The neglected terms in the Coulomb matrix (terms with more than two different orbital indices, i.e., the spin-flip exchange and correlated pair hopping terms) can significantly complicate the solution of the problem, in particular they give rise to a sign problem in LDA+DMFT calculations based on quantum Monte Carlo (QMC) solvers.

3.3 DFT+*U* method

The DFT+U method (also known as LDA+U or GGA+U) was the first attempt to combine many-body model Hamiltonian approaches and DFT [18, 26]. In the DFT+U method a Coulomb interaction term U is introduced as a phenomenological parameter to mimic correlation of the localized electrons. Since in the DFT+U method the Coulomb interaction term is treated on a static mean-field (unrestricted Hartree-Fock) level the true many-body effects are lost. Despite the different treatment of the U term in DFT+U the construction of the many-body Hamiltonian for the strongly correlated subspace is the same as the one we have discussed in the preceding section. To proceed let us consider the following simplified Hamiltonian, neglecting the Coulomb exchange term for a while

$$\hat{H}^{\text{DFT}+U} = \hat{H}^{\text{DFT}} + \frac{1}{2}U\sum_i \sum_{m\sigma\neq m'\sigma'} \hat{n}_{im\sigma}\hat{n}_{im'\sigma'} - \hat{H}^l_{\text{DC}}\,, \tag{35}$$

where the double counting term is given by the expression

$$\hat{H}^l_{\text{DC}} = \frac{1}{2}U\sum_i \sum_{m\sigma\neq m'\sigma'} \langle\hat{n}_{im\sigma}\rangle\langle\hat{n}_{im'\sigma'}\rangle\,, \tag{36}$$

which can be approximated by $\frac{1}{2}UN^l(N^l-1)$, where $N^l = N^l_\uparrow + N^l_\downarrow = \sum_{m\sigma}\langle\hat{n}_{im\sigma}\rangle$ is the number of localized electrons per site in the strongly correlated subspace. We will treat the Coulomb interaction on static mean-field Hartree level. To this end we first write the particle number operator as

$$\hat{n}_{im\sigma} = \langle\hat{n}_{im\sigma}\rangle + \delta\hat{n}_{im\sigma}\,, \tag{37}$$

where the fist part, i.e., $\langle\hat{n}_{im\sigma}\rangle$ corresponds to the expectation value and the second part $\delta\hat{n}_{im\sigma}$ contains quantum fluctuations. Then the Coulomb interaction term in the Hamiltonian takes the form

$$\begin{aligned}\frac{1}{2}U\sum_i \sum_{m\sigma\neq m'\sigma'} \hat{n}_{im\sigma}\hat{n}_{im'\sigma'} \;=\; & \frac{1}{2}U\sum_i \sum_{m\sigma\neq m'\sigma'} [\hat{n}_{im\sigma}\langle\hat{n}_{im'\sigma'}\rangle + \langle\hat{n}_{im\sigma}\rangle\hat{n}_{im'\sigma'} \\ & -\langle\hat{n}_{im\sigma}\rangle\langle\hat{n}_{im'\sigma'}\rangle + \cancel{\delta\hat{n}_{im\sigma}\delta\hat{n}_{im'\sigma'}}]\,,\end{aligned} \tag{38}$$

where the last term, i.e., correlation of fluctuations are neglected in mean-field approximation and the third term $\langle \hat{n}_{im\sigma} \rangle \langle \hat{n}_{im'\sigma'} \rangle$ is approximated by $\frac{1}{2} U N^l (N^l - 1)$. With these approximations the mean-field DFT+U Hamiltonian can be written as

$$\hat{H}^{\text{DFT}+U} = \hat{H}^{\text{DFT}} + \sum_{im\sigma} V_m^\sigma \hat{n}_{im\sigma} \, , \tag{39}$$

where $V_m^\sigma = U\big(\frac{1}{2} - \langle \hat{n}_{im\sigma} \rangle\big)$. The levels of the localized electrons in strongly correlated subspace are shifted $-U/2$ ($U/2$) if occupied (unoccupied). A total energy functional, which shifts the orbital energies is given by

$$E_{\text{DFT}+U}[n] = E_{\text{DFT}}[n] + \sum_i \left[\frac{1}{2} \sum_{m\sigma \neq m'\sigma'} \langle \hat{n}_{im\sigma} \rangle \langle \hat{n}_{im'\sigma'} \rangle - E_{\text{DC}} \right] , \tag{40}$$

where $E_{\text{DC}} = \frac{1}{2} N^l (N^l - 1)$ is the double counting Coulomb energy and $E_{\text{DFT}}[n]$ is the total energy obtained from the spin-polarized DFT calculation. The correction to the atomic orbital energies due to the Coulomb interaction U is given by

$$\epsilon_{im\sigma}^{\text{DFT}+U} = \frac{\partial E_{\text{DFT}+U}}{\partial \langle \hat{n}_{im\sigma} \rangle} = \epsilon_{im\sigma}^{\text{DFT}} + U\big(\frac{1}{2} - \langle \hat{n}_{im\sigma} \rangle\big) \, . \tag{41}$$

For a rotationally invariant Coulomb matrix [see Eq. (17)] the DFT+U total energy functional can be written as

$$\begin{aligned} E_{\text{DFT}+U}[n] \;=\; & E_{\text{DFT}}[n] + \frac{1}{2} \sum_{i\sigma} \sum_{mm'm''m'''} U_{m,m'';m',m'''} \langle \hat{n}_{imm'}^{\sigma} \rangle \langle \hat{n}_{im''m'''}^{-\sigma} \rangle \\ & + \frac{1}{2} \sum_{i\sigma} \sum_{mm'm''m'''} [U_{m,m'';m',m'''} - U_{m,m'';m''',m'}] \langle \hat{n}_{imm'}^{\sigma} \rangle \langle \hat{n}_{im''m'''}^{\sigma} \rangle \\ & - E_{\text{DC}} \, , \end{aligned} \tag{42}$$

where $\langle \hat{n}_{imm'}^{\sigma} \rangle$ is the density matrix and the double counting term E_{DC} in fully localized limit is given by

$$E_{\text{DC}}^l = \frac{1}{2} U_{av} N^l (N^l - 1) - \frac{1}{2} J_{av} \sum_\sigma N_\sigma^l (N_\sigma^l - 1) \, . \tag{43}$$

The corresponding one-electron mean-field DFT+U Hamiltonian in general form is

$$\hat{H}^{\text{DFT}+U} = \hat{H}^{\text{DFT}} + \sum_{imm'\sigma} V_{mm'}^\sigma c_{im\sigma}^\dagger c_{im'\sigma} \, , \tag{44}$$

where the one-electron potential including the double counting correction is defined by the expression

$$\begin{aligned} V_{mm'}^\sigma \;=\; & \sum_{im''m'''} U_{m,m'';m',m'''} \langle \hat{n}_{im''m'''}^{-\sigma} \rangle \\ & + \sum_{im''m'''} [U_{m,m'';m',m'''} - U_{m,m'';m''',m'}] \langle \hat{n}_{im''m'''}^{\sigma} \rangle \\ & - \left[U_{av}\big(N^l - \frac{1}{2}\big) - J_{av}\big(N_\sigma^l - \frac{1}{2}\big) \right] \delta_{m,m'} \, . \end{aligned} \tag{45}$$

Eqs. (42) and (44) are invariant with respect to orbital basis transformation. Note that the DFT+U method equations can be derived also in a variational way from the DFT+U functional as shown in Ref. [18]. An orbital dependent potential enters now the Kohn-Sham equation via

$$\big[-\frac{1}{2}\nabla^2 + v^{\sigma}_{\rm DFT}(\mathbf{r})\big]\psi^{\sigma}_n(\mathbf{r}) + \sum_{i,m,m'} V^{\sigma}_{mm'}\frac{\delta\langle\hat{n}^{\sigma}_{imm'}\rangle}{\delta\psi^{\sigma}_n(\mathbf{r})} = \epsilon_n\psi_n(\mathbf{r})\,. \tag{46}$$

Thus, a Hartree-Fock like potential term is introduced that acts on the orbitals of the strongly correlated subspace. Eq. (46) should be solved self-consistently, until both density matrix and density are converged. The results of the DFT+U calculations will depend to some extend on the choice of the basis set because such calculations are usually not performed using Wannier functions, a basis set, which is used for the calculation of Coulomb interaction parameters (see Section 4). They are typically based on the identification of the atomic spheres, a region where it is assumed that correlated electrons keep their atomic character and thus can be described by the atomic orbitals. Then the DFT+U corrections are determined by projection onto such atomic orbitals. However, if correlated electrons are well localized inside the atomic spheres the results do not depend on the choice of the basis set as confirmed by calculations based on different electronic structure methods. Apart from this, the results of the DFT+U calculations depend also on the treatment of the double counting problem. As discussed in preceding section there is no rigorous way to construct a $\hat{H}_{\rm DC}$ term. The usual approaches include i) the fully localized limit [18] as discussed above, ii) around mean-field method [27], and iii) interpolations between these two [28]. Therefore the DFT+U method can not be considered as a first principles method even if Coulomb interaction parameters are obtained from DFT calculations. Above we have presented the most complete formulation of the DFT+U method with orbital dependent electronic interactions. A simplified formulation is proposed by Dudarev *et al.*, in Ref. [29], which can be deduced from the full formulation by retaining only the lowest Slater integral F^0 (U) and neglecting the higher order ones, i.e., $F^2 = F^4 = F^6 = 0$ ($J = 0$).

Although DFT+U is a mean-field like approach it has been shown to be able to reproduce a correct band structure for the Mott insulators. The band gap in these materials originates from the strong on-site Coulomb interaction U, which splits the d (f) bands in a lower and upper Hubbard band, which is captured by the DFT+U method. Bands with d (f) character are shifted to lower energies if $\hat{n}^{\sigma}_{imm} > 0.5$ and higher energies if $\hat{n}^{\sigma}_{imm} < 0.5$. The value of U determines the magnitude of the energy shift. By increasing the Coulomb repulsion U the bands around Fermi energy are shifted either up or down so that the gap opens. As a simple example we consider ferromagnetic bcc Eu and employ the LDA+U method to correct the position of the 4f states. The calculated density of states and band structure is presented in Fig. 4. As seen the Coulomb correction U has almost no effect on the energy levels of the electrons of the weakly correlated subspace, i.e., s and p sates, but occupied (unoccupied) 4f states are shifted down (up) energies. Furthermore, an enhanced localization of the 4f states can be seen from the narrowing of the peaks in the density of states. Not only bcc Eu, but also other 4f metals and their compounds are studied with the DFT+U method. In particular, one of the problems of DFT-L(S)DA, the prediction of an antiferromagnetic ground state for hcp Gd, is resolved when the DFT+U method is employed [30].

More complicated physical phenomena, such as charge order, orbital order, etc. can be described by the DFT+U method. A collection of applications of the DFT+U method can be found in reference [26]. Although DFT+U method overcomes some major deficiencies of the DFT-LDA, it completely fails for strongly correlated systems, in which electrons simultaneously possess itinerant and localized character. In Green function language DFT+U method

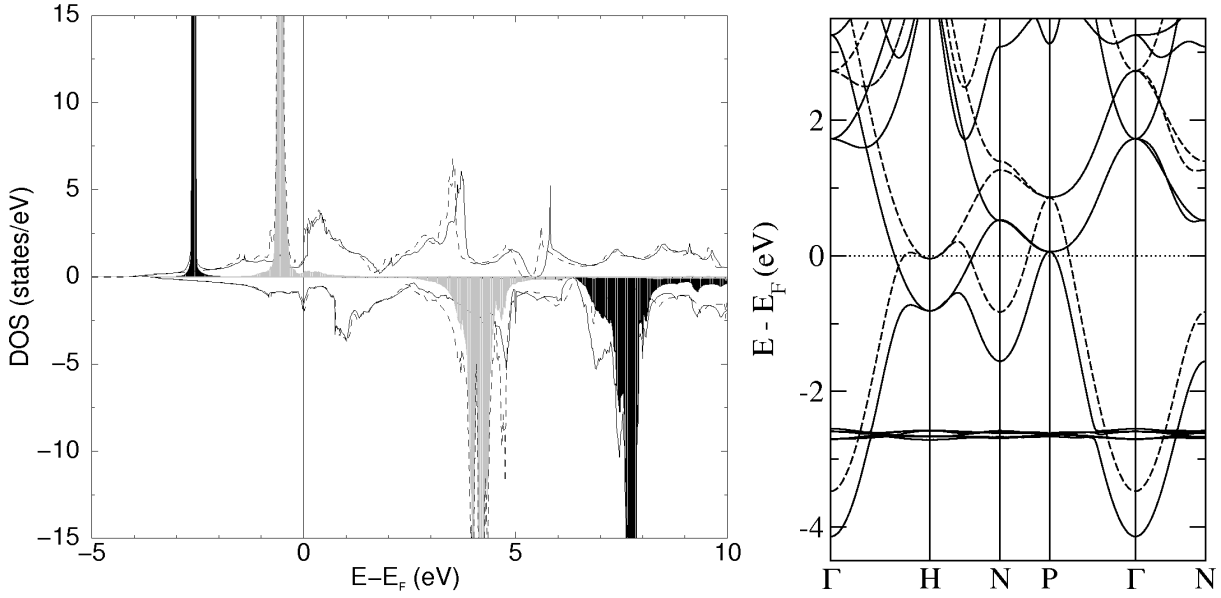


Fig. 4: *Left: Density of states (DOS) of ferromagnetic bcc Eu calculated with* LSDA *(dashed line) and the* LSDA+*U (full line) method. A Hubbard U value of about 6 eV is used to correct the position of the 4f states. The local partial DOS as obtained in* LSDA *is shown as grey shaded area, the* LSDA+*U result in black. Right: The band structure of bcc Eu obtained in the* LSDA+*U method. Majority (minority) spin stats is shown with full (broken) lines.*

possess a self energy, which is spin, orbital, and site dependent, but has no frequency dependence and thus can not capture the subtle correlation effects. For such materials one has to go beyond the one-electron picture and include correlations explicitly like in LDA+DMFT approach.

4 *Ab-initio* determination of the Coulomb interaction parameter *U*

The problem of calculating Coulomb interaction parameters for solids from first principles, has been addressed by several authors [31, 32, 33, 34, 35, 36, 37, 38], and a number of different approaches have been proposed. Among them, constrained local-density approximation (cLDA) [31, 32, 33], one of the earliest approach, is the most popular one. The cLDA is based on the fact that the energy of the system with increased or reduced electron number is in principle accessible within the DFT. The Hubbard U in cLDA is then defined as the derivative of the total energy with respect to the constrained occupation number in a given shell. A further improvement of this method was recently proposed [32]. The cLDA approach has been implemented in several electronic structure codes. However, cLDA is known to give unreasonably large U values for late transition metals and their compounds due to difficulties in compensating the self-screening error of the localized electrons [36]. Furthermore, the frequency dependence of the Coulomb interaction U is unattainable. A detailed discussion of the cLDA method can be found in reference [39]. On the other hand, the recently proposed constrained random-phase approximation (cRPA) [35, 36], though numerically much more demanding, does not suffer from these difficulties. It also allows to access individual Coulomb matrix elements, e.g., on-site,

off-site, intra-orbital, inter-orbital, and exchange as well as their frequency dependence. In this section we will describe the cRPA method.

4.1 Screened Coulomb interaction

The calculation of a Coulomb interaction between two electrons in vacuum is rather straightforward. However, the same can not be said for the interaction between two electrons in a solid, which is far from being trivial as the electronic polarization screens the Coulomb potential, giving rise to a renormalized repulsion strength. In solid state physics the concept of Coulomb screening is crucial to understand physical properties of materials, especially for metals. For instance, within Hartree-Fock approach, which does not take into account screening of the Coulomb interaction the density of states at the Fermi energy of metals vanishes [40]. When an electronic system is perturbed by an external static (or dynamic) potential the electrons will react to this potential in such a way that they minimize the total energy. If a positive charge is introduced to the electronic system, then the electrons will be attracted to this charge so as to reduce the total energy, i.e., the negative Coulomb potential energy will compensate the increase in the kinetic energy due to the localization of electrons around the positive test charge. As a consequence of the electron accumulation around the positive test charge the effective attraction between the positive charge and an electron far from it will be much weaker than the bare Coulomb interaction. In other words, Coulomb interaction will be screened by the motion of the other electrons. In the case of a negative test charge the other electrons will be repelled from it and a screening hole around the test charge will be created, which in a similar way screens the bare Coulomb interaction. For a general time dependent (dynamic) external field the screening will be also time dependent resulting in frequency dependent retarded interaction, which might become negative at finite frequencies.
Dynamically screened Coulomb interaction within the random-phase approximation (RPA) (see Lectrue A4 by C. Friedrich for a detailed discussion) is given by

$$\tilde{U}(\mathbf{r},\mathbf{r}';\omega) = \int d^3r'' \varepsilon^{-1}(\mathbf{r},\mathbf{r}'';\omega) v(\mathbf{r}'',\mathbf{r}') \,, \tag{47}$$

where $\varepsilon(\mathbf{r},\mathbf{r}'';\omega)$ is the dielectric function and $v(\mathbf{r}'',\mathbf{r}')$ being the bare Coulomb potential. The $\tilde{U}(\mathbf{r},\mathbf{r}';\omega)$ can be regarded as the effective potential at point $\mathbf{r}'$ induced by the electron at point $\mathbf{r}$. In RPA the dielectric function is expressed by

$$\varepsilon(\mathbf{r},\mathbf{r}';\omega) = \delta(\mathbf{r}-\mathbf{r}') - \int d^3r'' v(\mathbf{r},\mathbf{r}'') P(\mathbf{r}'',\mathbf{r}';\omega) \,, \tag{48}$$

where $P(\mathbf{r},\mathbf{r}';\omega)$ is the polarization function, which is given by

$$\begin{aligned} P(\mathbf{r},\mathbf{r}';\omega) &= \sum_{\sigma} \sum_{\mathbf{k},m}^{\text{occ}} \sum_{\mathbf{k}',m'}^{\text{unocc}} \varphi^{\sigma}_{\mathbf{k}m}(\mathbf{r}) \varphi^{\sigma *}_{\mathbf{k}'m'}(\mathbf{r}) \varphi^{\sigma *}_{\mathbf{k}m}(\mathbf{r}') \varphi^{\sigma}_{\mathbf{k}'m'}(\mathbf{r}') \\ &\quad \times \left[\frac{1}{\omega - \Delta^{\sigma}_{\mathbf{k}m,\mathbf{k}'m'}} - \frac{1}{\omega + \Delta^{\sigma}_{\mathbf{k}m,\mathbf{k}'m'}} \right] \end{aligned} \tag{49}$$

with $\Delta^{\sigma}_{\mathbf{k}m,\mathbf{k}'m'} = \epsilon^{\sigma}_{\mathbf{k}'m'} - \epsilon^{\sigma}_{\mathbf{k}m} - i\delta$, the Kohn-Sham eigenvalues $\epsilon^{\sigma}_{\mathbf{k}m}$, and a positive infinitesimal δ. $\varphi^{\sigma}_{\mathbf{k}m}(\mathbf{r})$ are single-particle Kohn-Sham states of spin σ and band index m. The σ runs over both

spin channels. In matrix notation the frequency dependent fully screened Coulomb interaction takes the form

$$\tilde{U}(\omega) = [1 - vP(\omega)]^{-1}v \, . \tag{50}$$

For metals and narrow band gap semiconductors the fully screened Coulomb interaction $\tilde{U}$ is almost independent of the bare interaction v, i.e., $v|P| \gg 1$ and thus $\tilde{U} \propto -1/P$. The polarization function P and as a consequence the $\tilde{U}$ is determined by the electronic structure of the screening electrons as will be discussed in Section 5.

4.2 Constrained random-phase approximation

The basic idea of the constrained random-phase approximation (cRPA) is to define an effective interaction U between the localized *d* (or *f*) electrons in the strongly correlated subspace by restricting the screening processes to the electrons of the weakly correlated subspace, which are not explicitly treated in the effective model Hamiltonian [36] (see also Section 3). As an example let us consider $SrVO_3$, whose band structure is shown in Fig. 1. We divide the one-particle Hilbert space into two parts, which are called l- and r-subspaces (see Fig. 3). The former contains the localized states of strongly correlated electrons, for which Coulomb interaction parameters are to be calculated, while the latter contains the states of weakly correlated electrons. The total polarization matrix P can now be separated into two parts

$$P = P_{\rm l} + P_{\rm r}, \tag{51}$$

where $P_{\rm l}$ includes only transitions between localized states, and $P_{\rm r}$ is the remainder. Then, the frequency-dependent effective Coulomb interaction between localized electrons in l-subspace is given by the matrix equation

$$U(\omega) = [1 - vP_{\rm r}(\omega)]^{-1}v \, . \tag{52}$$

When this effective interaction $U(\omega)$ is screened further by the electrons in the strongly correlated l-subspace one obtains fully screened Coulomb interaction

$$\tilde{U}(\omega) = [1 - U(\omega)P_{\rm l}(\omega)]^{-1}U(\omega) \, . \tag{53}$$

Note that in model Hamiltonian studies the $U(\omega)$ can be regarded as the bare Coulomb interaction for the strongly correlated l-subspace. The U is nonlocal in space and inherits a frequency dependence from $P_{\rm r}(\mathbf{r}, \mathbf{r}'; \omega)$. We consider matrix elements of U in the MLWF basis [see Eq. (12)]

$$U^{\sigma_1\sigma_2}_{i,n_1n_3;n_2n_4}(\omega) = \iint w^{\sigma_1 *}_{in_1}(\mathbf{r}) w^{\sigma_2 *}_{in_3}(\mathbf{r}') U(\mathbf{r}, \mathbf{r}'; \omega) w^{\sigma_2}_{in_4}(\mathbf{r}') w^{\sigma_1}_{in_2}(\mathbf{r}) \, d^3r \, d^3r' . \tag{54}$$

For the definition of the Hubbard U different conventions exist in the literature (for a detailed discussion see Ref. [41]). Historically, the Hubbard U was introduced as a Coulomb repulsion parameter between electrons in the single-orbital Hubbard-Kanamori-Gutzwiller model [42]. Note, that there is no Hund exchange J for a single orbital. For multi-orbital systems, the Hubbard U is defined as the static limit of the average intra-orbital and inter-orbital Coulomb matrix elements, i.e., Slater integral F^0 [see Eq. (30)]. Here, we follow the convention used in

Ref. [43] and denote the Hubbard U as $U_{\text{DFT+}U}$ and the Hund exchange interaction as J, which are given by

$$U_{\text{DFT+}U} = \frac{1}{L^2} \sum_{m,n} U^{\sigma_1\sigma_2}_{i,mn;mn}(\omega = 0) \,, \tag{55}$$

$$J = \frac{1}{L(L-1)} \sum_{m \neq n} U^{\sigma_1\sigma_2}_{i,mn;nm}(\omega = 0) \,, \tag{56}$$

where L is the number of localized orbitals. We note that for magnetic systems although the matrix elements of the Coulomb potential U are formally spin dependent due to the spin dependence of the MLWFs, this dependence is however negligible in practice. In DFT+U calculations the $U_{\text{DFT+}U}$ and J are used as the Hubbard U and Hund exchange J parameters, respectively, while for model Hamiltonian studies it is necessary to introduce the so-called Hubbard-Kanamori parametrization of the Coulomb matrix given by Eq. (54), in which for the same orbitals ($m = n$)

$$U = \frac{1}{L} \sum_{m} U^{\sigma_1\sigma_2}_{i,mm;mm}(\omega = 0) \tag{57}$$

and for different orbitals ($m \neq n$)

$$U' = \frac{1}{L(L-1)} \sum_{m \neq n} U^{\sigma_1\sigma_2}_{i,mn;mn}(\omega = 0) \,. \tag{58}$$

If the crystal field has a cubic symmetry, then the U' is given by $U' = U - 2J$. In this case, only two among U, U' and J are independent parameters. In multi-orbital systems, the Hund exchange J favors spin polarization. Similar to U, U', and J, we can also define the so-called fully screened $\tilde{U}$, $\tilde{U}'$, and $\tilde{J}$. Although the fully screened Coulomb interaction matrix elements are not used in model Hamiltonians, they provide an idea about the correlation strength of the considered electrons.

4.3 cRPA for entangled bands

Although, cRPA is a general approach, its application to materials with entangled bands is not straightforward. In these materials the localized *d* (or *f*) states that span the strongly correlated subspace mix with extended *s* and *p* states, and there is no unique identification of the *d*-*d* (or *f*-*f*) transitions for constructing P_{l}. Several procedures have been proposed in the literature to overcome this problem. Aryasetiawan *et al.* suggested to use an energy window or a range of band indices to define the l-subspace [36]. However, the results depend strongly on the chosen window or band indices. An alternative approach [37], in which the hybridization of the localized *d* (or *f*) states is switched off, is not burdened by additional parameters, but the U values turned out to be unphysically large for materials with strong *sp*-*d* mixing, e.g., early transition metals. Furthermore, the unphysical suppression of hybridization is unsatisfactory.
In a recent work [38], we propose a new parameter-free procedure where P_{l} is directly constructed from the definition of the l-subspace. The latter is spanned by a set of MLWFs. We now define P_{l} as the polarization function that is generated by all transitions that take place within the l-subspace. First we determine the probability of finding a *d* (or *f* electron) electron at state m with wave vector $\mathbf{k}$ within the l-subspace, which is given by

$$c^{\sigma}_{\mathbf{k}m} = \sum_{i,n} |T^{\sigma(\mathbf{k})}_{i,mn}|^2 \,, \tag{59}$$

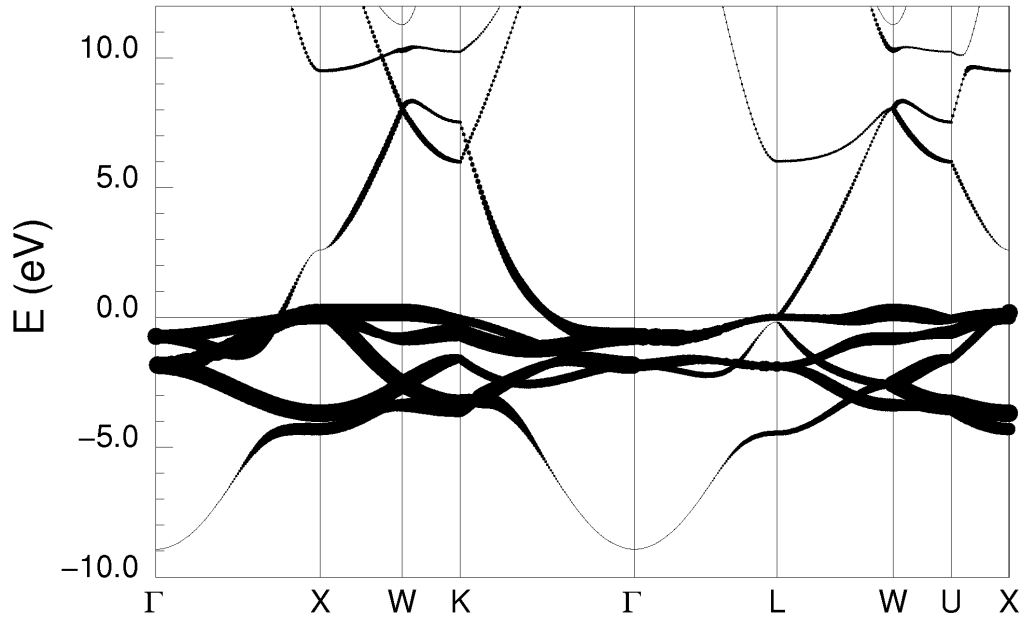


Fig. 5: *Band structure of non-magnetic fcc Ni. The probability factors $c^{\sigma}_{\mathbf{k}m}$ for 3d electrons (or 3d character of the bands) are shown by filled circles.*

where $T^{\sigma(\mathbf{k})}_{i,mn}$ is the unitary transformation matrix defined in Eq. (12). For entangled bands $0 < c^{\sigma}_{\mathbf{k}m} < 1$, while for disentangled (isolated) bands $c^{\sigma}_{\mathbf{k}m} = 1$. Then, the probability of a given transition between extended Bloch eigenstates $\varphi^{\sigma}_{\mathbf{k}m} \rightarrow \varphi^{\sigma}_{\mathbf{k}'m'}$ is given by

$$p^{\sigma}_{\mathbf{k}m\rightarrow\mathbf{k}'m'} = c^{\sigma}_{\mathbf{k}m} c^{\sigma}_{\mathbf{k}'m'} \; . \tag{60}$$

For isolated bands the factor $p^{\sigma}_{\mathbf{k}m\rightarrow\mathbf{k}'m'}$ is simply 1, but for the general case of entangled bands, one has $0 < p^{\sigma}_{\mathbf{k}m\rightarrow\mathbf{k}'m'} < 1$. As an example of entangled bands we consider non-magnetic fcc Ni and in Fig. 5 we present the band structure together with probability factors $c^{\sigma}_{\mathbf{k}m}$. As seen the flat bands around the Fermi energy are of almost 3*d* character i.e., $c^{\sigma}_{\mathbf{k}m} \simeq 1$, however a strong *s-d* mixing can be seen for the lower and higher bands leading to $0 < c^{\sigma}_{\mathbf{k}m} < 1$. Thus, the $P_{\rm l}$ is constructed from summing over all transitions in the Lehmann representation multiplied with these probabilities

$$\begin{aligned} P_{\rm l}(\mathbf{r},\mathbf{r}';\omega) \;\; = \;\; & \sum_{\sigma} \sum_{\mathbf{k},m}^{\rm occ} \sum_{\mathbf{k}',m'}^{\rm unocc} (p^{\sigma}_{\mathbf{k}m\rightarrow\mathbf{k}'m'})^2 \varphi^{\sigma}_{\mathbf{k}m}(\mathbf{r}) \varphi^{\sigma *}_{\mathbf{k}'m'}(\mathbf{r}) \\ & \times \varphi^{\sigma *}_{\mathbf{k}m}(\mathbf{r}') \varphi^{\sigma}_{\mathbf{k}'m'}(\mathbf{r}') \left[\frac{1}{\omega - \Delta^{\sigma}_{\mathbf{k}m,\mathbf{k}'m'}} - \frac{1}{\omega + \Delta^{\sigma}_{\mathbf{k}m,\mathbf{k}'m'}} \right] . \end{aligned} \tag{61}$$

In this way, the resulting effective interaction $U(\omega)$ only depends on the MLWFs that span the strongly correlated l-subspace and is basically independent of the used electronic structure method.

In a more formal approach, one can define $P_{\rm l}$ as the density correlation function

$$P_{\rm l}(\mathbf{r}t,\mathbf{r}'t') = -i\langle\Psi_0|\hat{T}[\hat{n}_{\rm l}(\mathbf{r}t),\hat{n}_{\rm l}(\mathbf{r}'t')]|\Psi_0\rangle \tag{62}$$

with the Kohn-Sham determinant Ψ_0, the time-ordering operator $\hat{T}$, and the Heisenberg density operator

$$\hat{n}(\mathbf{r}t) = \hat{n}_{\rm l}(\mathbf{r}t) + \hat{n}_{\rm r}(\mathbf{r}t) \tag{63}$$

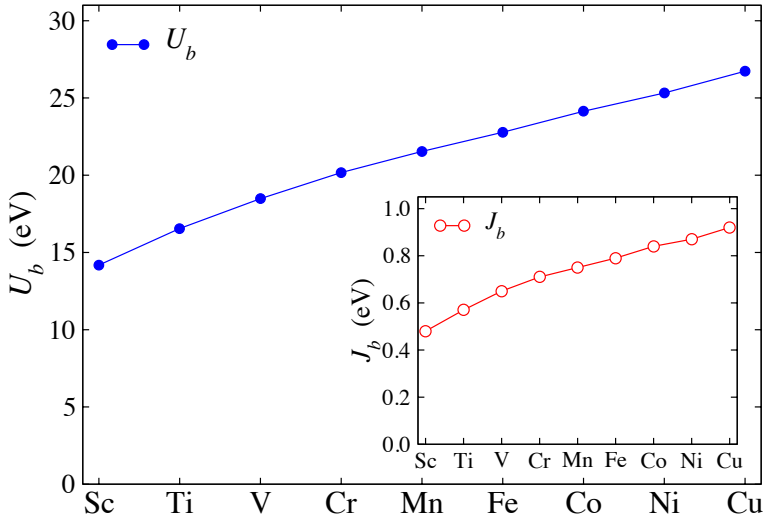


Fig. 6: *Average bare on-site direct Coulomb matrix elements between the d orbitals for 3*d *TMs. In the inset we show the results for exchange Coulomb matrix elements.*

decomposed according to the strongly correlated l-subspace and the rest. Without time-dependent external fields, P_{l} only depends on the time difference $t - t'$. A Fourier transformation then yields the Lehmann representation described above.

5 Applications

Employing the new cRPA method the Coulomb interaction parameters for different classes of materials are calculated [38, 44, 45, 46, 47, 48]. In this section, we will present Coulomb interaction parameters for 3*d* transition metals (TMs) and some prototype insulators. Both bulk phase and surfaces are considered. The ground-state calculations are carried out using the full-potential linearized augmented-plane-wave (FLAPW) method as implemented in the `FLEUR` code [50] with the GGA exchange-correlation potential as parameterized by Perdew *et al.* [51]. The MLWFs are constructed with the `Wannier90` code [49, 52]. The effective Coulomb interaction parameters are calculated within the recently developed cRPA method implemented in the `SPEX` code [53].

5.1 3*d* transition metals

As mentioned above cRPA calculation of Coulomb interaction parameters for systems with entangled bands like 3*d* TMs is not straightforward, since in these materials the correlated subspace formed by the 3*d* states can not be uniquely defined due to strong 3*d*-*sp* mixing. Previous cRPA studies of the Coulomb interaction U in 3*d* TMs have shown that the results strongly depend on the parameters used in the cRPA schemes [36, 37]. Below we will present results calculated with new cRPA method.

We start with the discussion of the *unscreened* (bare) Coulomb interaction. Figure 6 shows the average bare on-site direct (U_b) Coulomb matrix elements for the 3*d* TM series in the non-magnetic (NM) state. In the inset we show the results for the exchange (J_b) Coulomb matrix elements. Note that among the 3*d* series Fe, Co, and Ni are ferromagnetic (FM) while Cr orders antiferromagnetically. Also Mn is FM in the bcc structure with $a = 2.91$ Å. For these elements

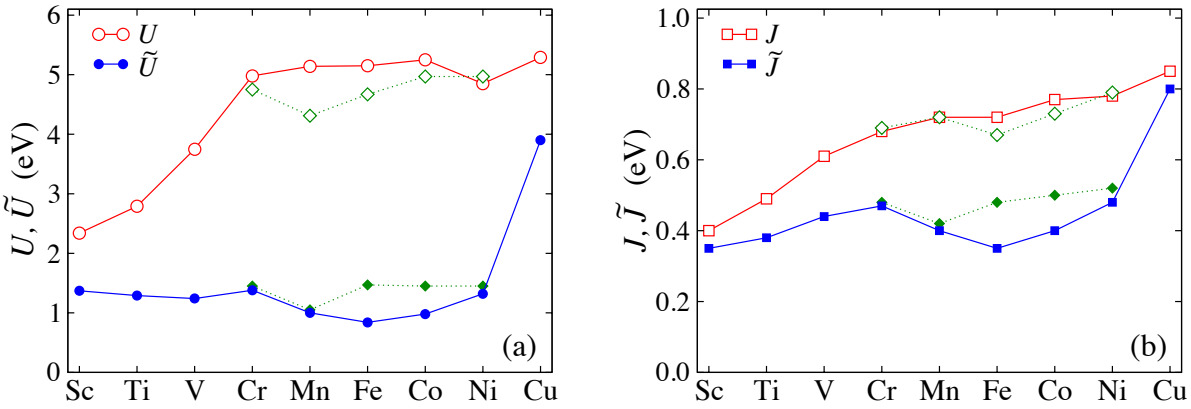


Fig. 7: *a) U and $\tilde{U}$ for the 3d TM series. With open and filled diamonds we show U and $\tilde{U}$ for the magnetic state of 3d TMs. b) The same for J and $\tilde{J}$.*

we find that matrix elements of the bare Coulomb potential for magnetic and NM states are nearly identical. As seen both U_b and J_b increase monotonically with the d electron number. This can be explained by the fact that, as one moves from the left to the right within one row of the periodic table, the nuclear charge increases and causes the d-wave functions to contract, which gives rise to the observed trend for U_b and J_b.

Efficient *sp* screening in 3*d* TMs significantly reduces the bare Coulomb interaction U_b. Calculated U and J parameters for the NM state of the 3*d* TMs are presented in Fig. 7. Results for the magnetic states are also included. For comparison, matrix elements of the fully screened Coulomb interaction ($\tilde{U}$, $\tilde{J}$) are given. As seen in Fig. 7, in contrast to the bare direct Coulomb interaction U_b, the partially screened U shows a non-monotonic behavior, i.e., it increases from the early 3*d* TMs and reaches a plateau-like behavior around half-filling, whereas $\tilde{U}$ is almost constant across the 3*d* TM series, except for the elements with completely filled d shells like Cu. This behavior of the U reflects a substantial contribution of the *d*-*d* transitions to the fully screened Coulomb interaction $\tilde{U}$, especially around half-filling. In difference to the bare Coulomb U_b, the partially screened U parameter is very sensitive to the d electron number and d orbital filing. In metals we are in the strong coupling limit, $v|P_{\rm r}| \gg 1$, and thus $U \propto -1/P_{\rm r}$. Since $P_{\rm r}$ depends mostly on the electronic structure of the screening electrons. This explains the observed trend for U values. Furthermore, by constrained NM and proper spin-polarized treatments of the magnetic elements, we show that spin polarization has a strong influence on U and $\tilde{U}$. Our calculated U parameters for the 3*d* series are in good agreement with recent cRPA studies of Miyake *et al.* [37] for late TMs as well as cLDA calculations of Nakamura *et al.* [33] for early TMs.

In contrast to U and U', renormalization of the J is rather small, i.e., J is close to the atomic value J_b. However, the *d*-*d* transitions substantially reduce $\tilde{J}$, especially for late 3*d* TMs. It should also be noted that while the bare U_b has a long-range behavior, the U shows much faster damping. The calculated nearest-neighbor U values lie between 0.1 and 0.4 eV being maximal for TMs with half-filled d bands.

5.2 Effective Coulomb interaction at surfaces of metals and insulators

The effective on-site Coulomb interaction U between localized electrons at surfaces of solids is expected to be enhanced since the effective screening volume of the surface is reduced with

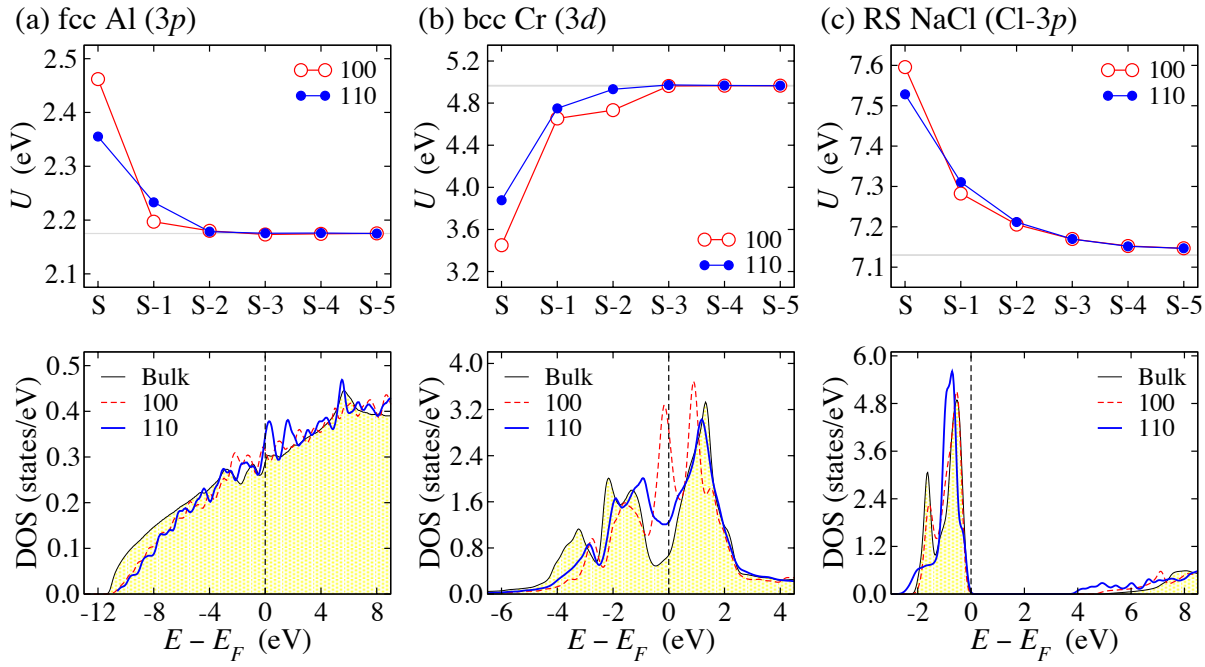


Fig. 8: *(a) Upper panel: layer dependence of the U parameter for fcc Al. The* S *denotes the surface layer in the slab model. Lower panel: Total density of states (DOS) for the (100) and (110) surface of fcc Al. For comparison, the bulk DOS (shaded area) is included; (b) and (c) the same for bcc Cr and rock-salt NaCl.*

respect to the bulk . As a consequence, the electron polarization decreases at the surface, which reduces the effect of screening and gives rise to a larger U value. Neither experimental nor theoretical works have been reported so far that would address the strength of the surface U parameter explicitly. However, a large number of phenomena observed in solids indicates an enhancement of the U at surfaces. For instance, the metal-insulator transition at the surface of correlated materials [54], the appearance of magnetism at the surface of paramagnetic transition metals [55], and the enhanced exchange splitting at the surface of 3*d* ferromagnets [56] have been attributed to an increase of the correlation strength U/W. In the theoretical description of surfaces, the U is usually assumed to be unchanged [57, 58, 59] so that the enhancement of correlation at the surface (S) with respect to the bulk (B) is provided by the effective band narrowing, i.e., $W_{\mathrm{S}} < W_{\mathrm{B}}$. In principle, depending on the relative values of the surface U and the bandwidth W, the correlation strength can further increase or decrease even below the bulk value. However, the latter case is considered to be unlikely because it is believed that U always increases at surfaces. In Ref. [46] we have shown by means of first-principles calculations that this is not always the case. It decreases at many TM surfaces and insulator surfaces with pronounced surface states, as a result of additional screening channels that open up due to surface-related changes in the electronic structure.

We model the metal and insulator surfaces with slabs of 11 atomic layers. Such slabs form a superlattice with 20 Å of vacuum separating them, with each slab possessing two (100) or (110) symmetric surfaces. Additional modifications, such as surface reconstruction and surface relaxation are not taken into account. We consider the 3*d* TMs in their respective ground-state crystal structures except Sc, Ti, and Co (Mn), which are treated in the fcc (bcc) structure

In Fig. 8 we present the calculated layer dependence of the Coulomb interaction parameter U

Table 1: *Bulk and surface U and J values for simple metals and insulators. The corresponding orbitals for which the U and J are calculated are given in parenthesis.*

	Bulk		100 Surface		110 Surface	
	U_{B}	J_{B}	U_{S}	J_{S}	U_{S}	J_{S}
Na [3*s*]	1.39		1.50		1.47	
Al [3*p*]	2.18	0.36	2.46	0.39	2.36	0.38
MgO [O-2*p*]	7.10	0.63	7.23	0.62	6.38	0.59
NaCl [Cl-3*p*]	7.13	0.55	7.60	0.54	7.53	0.55
$SrTiO_3$ [Ti-3*d*]	3.34	0.37	3.62	0.39		
$SrTiO_3$ [O-2*p*]	4.42	0.56	4.79	0.55		

for three different systems: (i) the simple metal fcc Al, (ii) the transition metal bcc Cr, and (iii) the insulator NaCl in the rock-salt structure. As seen in cases (i) and (iii), the U increases from the middle layer, where it is close to the bulk value, to the surface layer as expected. However, we find an unexpected behavior in the case of bcc Cr, where U is substantially reduced at the surface. This reduction is 30% for the open (100) and 20% for the (110) surface. For simplicity – the magnetism of bcc Cr is quite complicated – we only consider the non-magnetic state here. Furthermore, the layer dependence of U is quite different in metals and insulators. Because of the short screening length in metals, the U quickly assumes the bulk value in the former, as we go from the surface toward the middle of the slab, while in the latter the layer-by-layer convergence to the bulk value is much slower. On the other hand, the surface J values, which are listed in Table 1 for various materials, only differ slightly from the corresponding bulk values.

It has been shown that the surface electronic polarization and, as a consequence, the Coulomb interaction U is determined by two competing effects: (i) the so-called dimensionality effect, which is due to the reduced coordination number and hence the decrease of the effective screening volume at the surface region. From the point of view of classical electrostatics, this effect reduces the electronic polarization at the surface leading to larger U values. (ii) Electronic structure effects, i.e., the appearance of surface states and the effective band narrowing. This second effect gives rise to an enhancement of the electronic polarization and, hence, to a decrease of U. Depending on the strength of the two competing effects, the effective Coulomb interaction at the surface can be enhanced as well as reduced with respect to the bulk value. Qualitative information on the influence of the surface electronic structure on the interaction parameter U, leading to the second effect, can be deduced from the polarization function given in Eq. (49). For zero frequency ($\omega = 0$) the main contribution to the polarization function comes from the states around the Fermi energy. The smaller the energy difference between occupied and unoccupied states the larger the contribution. The effective band narrowing in TMs tends to reduce $\Delta^{\mathrm{S}}_{mm'}$ with respect to $\Delta^{\mathrm{B}}_{mm'}$, which has the effect of increasing the polarization. Additionally, the presence of surface states close to the Fermi level at the TM bcc (100) surface makes $\Delta^{\mathrm{S}}_{mm'}$ effectively smaller resulting in a more efficient electronic polarization and, as a result, in substantially reduced U and $\tilde{U}$ values as shown in Fig. 9 for the 3*d* series. As seen in Fig. 8, the density of states (DOS) of the (100) surface of bcc Cr exhibits a pronounced peak that is caused by a surface state just below the Fermi level, which is also found in previous first-principles calculations [60]. This surface state contributes substantially to the polarization function and is mainly responsible for the stronger reduction of the U value compared to the corresponding

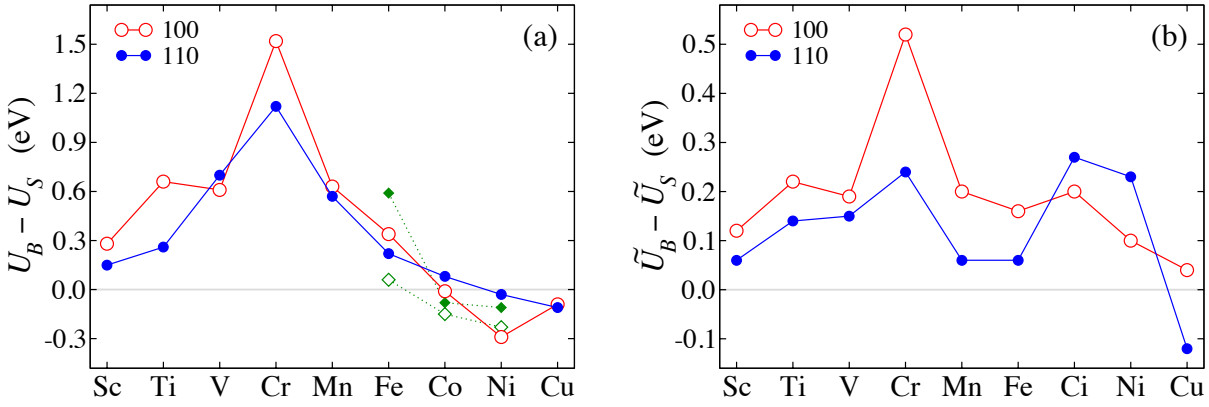


Fig. 9: *(a) The difference between the bulk and surface U values. Open [filled] diamonds show $U_\mathrm{B} - U_\mathrm{S}$ for the (100) [(110)] surface of ferromagnetic Fe, Co, and Ni. (b) The same as (b) for the fully screened Coulomb interaction $\tilde{U}$.*

value for the (110) surface, where such a peak in the DOS is missing. On the other hand, in simple metals the surface electronic structure turns out to be very similar to that of the bulk so that the dimensionality effect (i) wins over the electron-structure effect (ii), giving rise to enhanced Coulomb interaction parameters.

For most other 3*d* metals, the interaction parameter U is reduced at the surface, too, but less than in the case of bcc Cr, and for the late TMs the surface U exceeds that of the bulk U. The U and $\tilde{U}$ values for the bulk phase of 3*d* TMs are already presented in Fig. 7 The difference between the bulk and surface U and $\tilde{U}$ values, i.e., $U_\mathrm{B} - U_\mathrm{S}$ and $\tilde{U}_\mathrm{B} - \tilde{U}_\mathrm{S}$, is presented in Fig. 9(a) and (b) for the (100) and (110) surface. As seen, from Sc to Fe the U is reduced at both surfaces, and Co is at the border, in which the U assumes similar values in the bulk (see Fig. 7) and at the surface. Only at the Ni and Cu surface the U is slightly larger. As for the difference of the fully screened Coulomb interaction, i.e., $\tilde{U}_\mathrm{B} - \tilde{U}_\mathrm{S}$, we obtain a qualitatively similar behavior, but the relative reduction of $\tilde{U}$ with respect to the bulk value is significantly larger. This is attributed to the fact that, in contrast to the partially screened U, screening effects stemming from $3d \rightarrow 3d$ transitions contribute to the effective interaction, too. At the surface these virtual transitions take place within the surface states, which leads to very effective screening effects that give rise to the observed reduction of the $\tilde{U}$. It is important to point out that the variations of U seen in Fig. 9 are not caused by different spreads of the Wannier function across the series. We note that the U values depend only little on the bare Coulomb interaction because in metals we are in the strong coupling limit, i.e., $v|P_1| \gg 1$, and thus $U \simeq -P_1^{-1}$. As for the surface U of the ferromagnets Fe, Co, and Ni the same discussion holds.

In contrast to metals, the U at insulator surfaces is much more strongly affected by the presence of surface states than the band narrowing. For example, the NaCl (100) surface does not exhibit any surface states [61], and the U is enhanced. The slight changes in the electronic structure of the (110) surface (the gap is smaller, see Fig. 8) reduces the surface U only very little so that it is still larger than the bulk value. The situation is similar for the $SrTiO_3$ surface, for which the results are presented in Table 1. In MgO, on the other hand, both (100) and (110) surfaces exhibit surface states, which effectively reduce the band gap. For the (100) surface this gap reduction is around 0.6 eV [63], while for the (110) surface we obtain a much stronger reduction, the surface states lie 3.3 eV below the conduction-band minimum, i.e., close to the middle of the bulk band gap, which amounts to 4.97 eV. The presence of the surface state strongly affects the

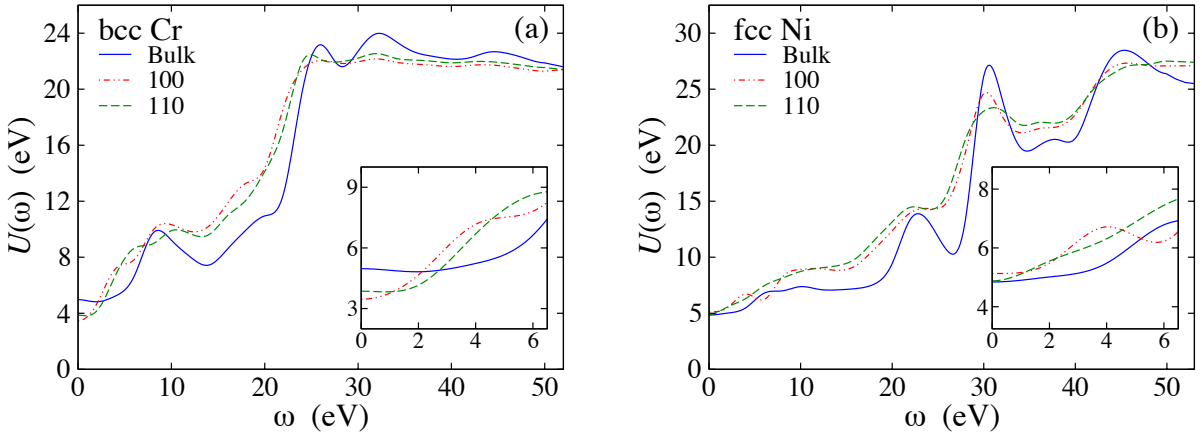


Fig. 10: *(a) Frequency dependence of the bulk and surface Coulomb interaction parameter U for bcc Cr. In the inset we expand the low frequency region. (b) The same for fcc Ni.*

screening properties with the consequence that the U parameter is considerably reduced at the (110) surface, while it remains slightly above the bulk value in the case of the (100) surface.
Finally we discuss the frequency dependence of the surface $U(\omega)$ considering the non-magnetic state of bcc Cr and fcc Ni. The obtained results are presented in Fig. 10 and compared with the bulk calculations. We see that, apart from the substantial reduction of the static U at the Cr surface, its frequency dependence is quite different from that of the bulk: $U_{\rm S}(\omega)$ increases monotonically with frequency, and at about 2 eV it crosses the $U_{\rm B}(\omega)$ curve towards larger values, while $U_{\rm B}(\omega)$ stays almost constant between 0 and 4 eV. For frequencies up to the plasmon frequency, $U_{\rm S}(\omega)$ is mostly larger than $U_{\rm B}(\omega)$. This behavior is not surprising because in the TMs the 3*d* states close to the Fermi level are much more affected by the presence of the surface than the states at higher energies. Hence, with increasing frequency the polarization function becomes less sensitive to subtleties of the surface electronic structure, and at high frequencies the surface $U(\omega)$ tends to be larger than the bulk value due to the dimensionality effect.

References

[1] W. Kohn, Rev. Mod. Phys. **71**, 1253 (1998).

[2] U. von Barth, Physica Scripta **109**, 9 (2004).

[3] P. Hohenberg and W. Kohn. Phys. Rev. B **76**, 6062 (1964).

[4] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[5] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[6] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, 1994).

[7] U. von Barth and L. Hedin J. Phys. C: Solid State Phys. **5**, 1629 (1972).

[8] I. V. Solovyev, J. Phys.: Condens. Matt. **20**, 293201 (2008).

[9] V. Anisimov and Y. Izyumov, *Electronic Structure of Strongly Correlated Materials* (Springer-Verlag Berlin, 2010).

[10] P. Fulde, *Electron Correlations in Molecules and Solids* (Springer, 1995).

[11] N. Hamada, H. Sawada, K. Terakura, in *Spectroscopy of Mott insulators and correlation metals*, ed. by A. Fujimori, Y. Tokura (Springer-Verlag, Berlin, 1995).

[12] F. Aryasetiawan and O. Gunnarsson, Rep. Prog. Phys. **61**, 237 (1998).

[13] W. Pickett, Rev. Mod. Phys. **61**, 433 (1989).

[14] A. Yanase, K. Siratori, J. Phys. Soc. Jpn. **53**, 312 (1984).

[15] Z. Zhang, S. Satpathy, Phys. Rev. B **44**, 13319 (1991).

[16] A. Georges, G. Kotliar, W. Krauth and M. Rozenberg, Rev. Mod. Phys. **68**, 13 (1996).

[17] V. I. Anisimov, A. I. Poteryaev, M. A. Korotin, A. O. Anokhin, and G. Kotliar, J. Phys.: Condens. Matter **9**, 7359 (1997).

[18] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).

[19] N. Marzari and D. Vanderbilt, Phys. Rev. B **56**, 12847 (1997).

[20] I. Souza, N. Marzari, and D. Vanderbilt, Phys. Rev. B **65**, 035109 (2001).

[21] J. F. Janak, Phys Rev B **18**, 7165 (1978).

[22] J. C. Slater, *Quantum theory of atomic structure*, Vol. 1 (McGraw-Hill, New York, 1960).

[23] B. Judd, in *Operator techniques in atomic spectroscopy*, (McGraw-Hill, New York, 1963).

[24] F. de Groot, J. Fuggle, B. Thole, and G. Sawatzky, Phys. Rev. B **42**, 5459 (1990).

[25] T. Fujiwara and M. Korotin, Phys. Rev. B **59**, 9903 (1999).

[26] V. I. Anisimov, F. Aryasetiawan and A. I. Lichtenstein, J. Phys.: Condens. Matter, **9**, 767 (1997).

[27] M. T. Czyżyk and G. A. Sawatzky, Phys. Rev. B **49**, 14211 (1994).

[28] F. Bultmark, F. Cricchio, O. Grånäs, and L. Nordström Phys. Rev. B **80**, 035121 (2009).

[29] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, Phys. Rev. B **57**, 1505 (1998).

[30] Ph. Kurz, G. Bihlmayer, and S. Blügel, J. Phys.: Condens. Matter **14**, 6353 (2002).

[31] P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai, Phys. Rev. Lett. **53**, 2512 (1984); V. I. Anisimov and O. Gunnarsson, Phys. Rev. B **43**, 7570 (1991).

[32] M. Cococcioni and S. de Gironcoli, Phys. Rev. B **71**, 035105 (2005).

[33] K. Nakamura, R. Arita, Y. Yoshimoto, and S. Tsuneyuki, Phys. Rev. B **74**, 235113 (2006).

[34] T. Kotani, J. Phys.: Condens. Matter **12**, 2413 (2000).

[35] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, Phys. Rev. B **70**, 195104 (2004).

[36] F. Aryasetiawan, K. Karlsson, O. Jepsen, and U. Schönberger, Phys. Rev. B **74**, 125106 (2006);

[37] T. Miyake, F. Aryasetiawan, and M. Imada Phys. Rev. B **80**, 155134 (2009).

[38] E. Şaşıoğlu, C. Friedrich, and S. Blügel, Phys. Rev. B **83**, 121101(R) (2011).

[39] B. Himmetoglu, A. Floris, S. de Gironcoli, M. Cococcioni, Int. J. Q. Chem. **114**, 14 (2014).

[40] N. W. Aschcroft and N. D. Mermin, *Solid State Physics* (Brookes Cole, 1976).

[41] L. Vaugier, H. Jiang, and S. Biermann, Phys. Rev. B **86**, 165105 (2012).

[42] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963); J. Kanamori, Prog. Theor. Phys. **30**, 275 (1963); M. C. Gutzwiller, Phys. Rev. Lett. **10**, 159 (1963).

[43] V. I. Anisimov, I. V. Solovyev, M. A. Korotin, M. T. Czyzyk, and G. A. Sawatzky, Phys. Rev. B **48**, 16929 (1993).

[44] E. Şaşıoğlu, I. Galanakis, C. Friedrich, and S. Blügel, Phys. Rev. B **88**, 134402 (2013).

[45] T. O. Wehling, E. Şaşıoğlu, C. Friedrich, A. I. Lichtenstein, M. I. Katsnelson, and S. Blügel, Phys. Rev. Lett. **106**, 236805 (2011).

[46] E. Şaşıoğlu, C. Friedrich, and S. Blügel, Phys. Rev. Lett. **109**, 146401 (2012).

[47] H. Sims, W. H. Butler, M. Richter, K. Koepernik, E. Şaşıoğlu, C. Friedrich, and S. Blügel, Phys. Rev. B **86**, 174422 (2012).

[48] H. Sims, E. R. Ylvisaker, E. E. Şaşıoğlu, C. Friedrich, S. Blügel, and W. E. Pickett, Phys. Rev. B **87**, 195120 (2013).

[49] F. Freimuth, Y. Mokrousov, D. Wortmann, S. Heinze, and S. Blügel, Phys. Rev. B **78**, 035120 (2008).

[50] http://www.flapw.de

[51] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[52] A. A. Mostofi, J. R. Yates, Y.-S. Lee, I. Souza, D. Vanderbilt, and N. Marzari, Comput. Phys. Commun. **178**, 685 (2008).

[53] C. Friedrich, S. Blügel and A. Schindlmayr, Phys. Rev. B. **81**, 125102 (2010).

[54] S.-K. Mo, J. D. Denlinger, H.-D. Kim, J.-H. Park, J. W. Allen, A. Sekiyama, A. Yamasaki, K. Kadono, S. Suga, Y. Saitoh, T. Muro, P. Metcalf, G. Keller, K. Held, V. Eyert, V. I. Anisimov, and D. Vollhardt, Phys. Rev. Lett. **90**, 186403 (2003).

[55] A. Goldoni, A. Baraldi, G. Comelli, S. Lizzit, and G. Paolucci, Phys. Rev. Lett. **82**, 3156 (1999).

[56] N. Kamakura, Y. Takata, T. Tokushima, Y. Harada, A. Chainani, K. Kobayashi and S. Shin, Europhys. Lett. **67**, 240 (2004).

[57] M. Potthoff and W. Nolting, Phys. Rev. B **59**, 2549 (1999)

[58] A. Liebsch, Phys. Rev. Lett. **90**, 096401 (2003).

[59] A. Grechnev, I. Di Marco, M. I. Katsnelson, A. I. Lichtenstein, J.Wills, and O.Eriksson, Phys. Rev. B **76**, 035107 (2007).

[60] O. Yu. Kolesnychenko, G. M. Heijnen, A. K. Zhuravlev, R. de Kort, M. I. Katsnelson, A. I. Lichtenstein, and H. van Kempen, Phys. Rev. B **72**, 085456 (2005).

[61] B. Li, A. Michaelides, and M. Scheffler, Phys. Rev. B **72**, 085408 (2005).

[62] Y. Adachi, S. Kohiki, K. Wagatsuma, and M. Oku, J. Appl. Phys. **84**, 2123 (1998).

[63] U. Schönberger and F. Aryasetiawan, Phys. Rev. B **52**, 8788 (1995).

A 10 Relativistic effects in solids [1]

G. Bihlmayer
Peter Grünberg Institut and
Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Since the title of this contribution is very general and it is obviously not possible to discuss in two hours the field of "relativistic effects" in full depth, I will present a selection of what is interesting and useful in the context of "computing solids". Most prominently it is devoted to spin-orbit coupling (SOC) as it appears in different phenomena, in magnetic and non-magnetic materials. But we will encounter also other "relativistic terms" that can be derived from the (single-particle) Dirac equation and some "relativistic physics" in the context of graphene and topological insulators.

Typically, as a student we encounter the term "spin-orbit coupling" first in context of atomic physics: After we learn how to solve the Schödinger equation for a hydrogenic atom, where all the energy levels can be labeled with a single (principal) quantum number, we are confronted with the fact that these levels have additional structure caused by a relativistic effect. This effect, SOC, splits the p-levels into $p_{\frac{1}{2}}$ and $p_{\frac{3}{2}}$, causing the fine-structure of the hydrogen atom, one of the first evidences for a property called "spin" of the electron. Spin is an attribute of the electron that emerges naturally from an relativistic single particle theory, the Dirac equation, and SOC is also naturally described in this framework.

Not only historically the spin and spin-orbit coupling are of relevance, we encounter these concepts nowadays in information technology, in particular for storage and manipulation of data encoded in the spin or the electron. The control of individual spins is at the heart of this process and for the success of what is called "magnetoelectronics" or "spintronics" it is mandatory to understand the underlying physical principles in the solid state. Like the charge state of a device can not only be used to store, but also to process data, it can be envisaged that here also the spin degrees of freedom are utilized for these purposes. In contrast to the charge, which is scalar, the spin is a vector quantity and, therefore, not only the magnitude but also the direction of a spin can be manipulated. Thus, at least in principle, spin-based electronics (spintronics) allows for new concepts which have no analogon in charge based devices.

But how can we control or manipulate the spin? On the level of a non-relativistic (Schrödinger) theory spin appears just as an additional quantum number, a label to introduce the Pauli exclusion principle for indistinguishable particles like electrons. On the other hand, we know that there is a angular momentum connected with the electrons spin, leading to a magnetic moment that can couple to external magnetic fields (Stern-Gerlach experiment). But for practical applications on the nanoscale magnetic fields are cumbersome as they are hard to confine. So, how can we connected the electrons spin with a non-magnetic "outside world", e.g. the lattice of a crystal? On a single-particle level it turns out that SOC provides the first term that couples the spin direction to potential variations of the real space, i.e. local electrical fields that arise naturally or are controlled externally.

Finally, spin-orbit coupling has recently gained a lot of attention in the context of topological insulators (TIs). There, SOC acts as a kind of 'built-in magnetic field' of a non-magnetic solid, that plays a role similar to the external magnetic field in the quantum Hall effect (QHE). Like in the QHE it creates a state of matter that is characterized by a topological number and unique signatures in electronic transport with precise quantization. However, this analogy is only true if just one type of spin is considered, on particles with the other spin SOC acts as magnetic field of opposite direction. Nevertheless, the effects do not cancel but lead to a new phenomenon, the quantum spin-Hall effect in TIs. Like in the QHE, "topologically protected" transport arises and holds promise for future spintronic applications or quantum computing.

Although this list of SOC-related phenomena, that will be covered in this lecture, is already

quite long, there are aspects of spin-orbit coupling that will not be included here: Several important scattering effects that gain their spin-selectivity from SOC will be treated in the lecture on scattering theory by Ph. Mavropoulos (A5). More insight in SOC-related topics in magnetic systems can be obtained from the lecture on complex magnetism by S. Blügel (C4), here we will just cover the basic interactions arising from SOC. For this propose, we start with a discussion of the Dirac equation and one-electron atoms, before we go over to the nonmagnetic solids and relativistic effects therein and finally study the influence of exchange interactions on these phenomena.

2 Spin-orbit coupling in atoms

2.1 The Dirac equation

The relativistic theory for an electron (of charge $-e$ and mass m) in an external scalar potential V and a vector potential $\mathbf{A}$ is given by the Dirac equation

$$\hat{H}\Psi = +i\hbar\frac{\partial}{\partial t}\Psi = E'\Psi; \qquad \hat{H} = -eV(\mathbf{r}) + \beta mc^2 + \boldsymbol{\alpha}\cdot(c\mathbf{p} + e\mathbf{A}(\mathbf{r})). \tag{1}$$

Here, $\boldsymbol{\alpha}$ is a vector of 4×4 matrices that can be written in terms of the Pauli spin-matrices, $\boldsymbol{\sigma}$, while β is a matrix of same rank, expressible in terms of the 2×2 unit matrix I_2:

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \beta = \begin{pmatrix} I_2 & 0 \\ 0 & -I_2 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2}$$

The Hamiltonian acts on a four-component (bi-spinor) wavefunction Ψ that describes particles with the total energy E' (the rest mass energy mc^2 included). Discussions of the Dirac equation are available in most textbooks, here we follow Ref. [1]. To extract a 'spin-orbit coupling' from this theory, it is useful to make contact with the Schrödinger equation again. The four-component wavefunction can be written as a 2-vector of the so-called large and small components, ψ and χ. For these components the Dirac equation can be written as

$$(E' - mc^2 + eV(\mathbf{r}))\psi = \boldsymbol{\sigma}\cdot(c\mathbf{p} + e\mathbf{A}(\mathbf{r}))\chi \tag{3}$$

$$(E' + mc^2 + eV(\mathbf{r}))\chi = \boldsymbol{\sigma}\cdot(c\mathbf{p} + e\mathbf{A}(\mathbf{r}))\psi\,. \tag{4}$$

In the non-relativistic limit, we can assume that the energy is close to the rest mass energy and neglect $eV(\mathbf{r}) << E' + mc^2 \approx 2mc^2$. Within this assumption, we can use equation (4) to approximate χ and insert into equation (3). Writing $E = E' - mc^2$ these equations reduce then to a single equation for the large component:

$$\left[E + eV(\mathbf{r}) - \frac{1}{2m}\left(\mathbf{p}(\mathbf{r}) + \frac{e}{c}\mathbf{A}(\mathbf{r})\right)^2\right]\psi = 0 \tag{5}$$

which is a minimal extension of the Schrödinger equation for a particle in a vector potential. Substituting the expression for the momentum operator, $\mathbf{p}(\mathbf{r}) = -i\hbar\boldsymbol{\nabla}$ and neglecting $\mathbf{A}(\mathbf{r})$ gives back the usual presentation of the Schrödinger equation. We see that the large component is, in this approximation, equivalent to the wavefunction of the Schrödinger equation. The Pauli spin-matrices dropped out from the equation and only the fact that ψ is a spinor reminds us of the electron's spin.

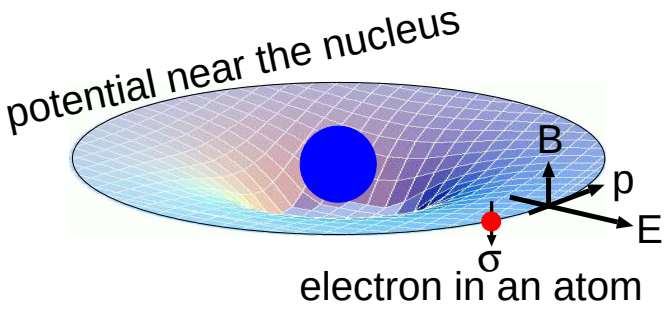


Fig. 1: *Schematic illustration of an electron moving in the potential gradient near a nucleus. The electric field* $\mathbf{E}$ *resulting from the potential gradient is Lorentz-transformed into a* $\mathbf{B}$ *field by the orbital motion of the electron. The electron's spin,* $\boldsymbol{\sigma}$ *couples to the resulting* $\mathbf{B}$ *field.*

To obtain terms that explicitly involve the spin-matrices, we have to refine our approximation to the Dirac equation. In principle, this can be done by obtaining with the above equations an approximation for χ substituting in (3) and retaining terms up to order $(v/c)^2$. It is possible to formulate an equation (sometimes termed Pauli equation, see [1]) for the large component only:

$$\begin{aligned}\Big[E + eV(\mathbf{r}) - \frac{1}{2m}\Big(\mathbf{p}(\mathbf{r}) + \frac{e}{c}\mathbf{A}(\mathbf{r})\Big)^2 + \frac{1}{2mc^2}\left(E + eV(\mathbf{r})\right)^2 \quad + \\ i\frac{e\hbar}{(2mc)^2}\mathbf{E}(\mathbf{r})\cdot\mathbf{p} - \frac{e\hbar}{(2mc)^2}\boldsymbol{\sigma}\cdot(\mathbf{E}(\mathbf{r})\times\mathbf{p}) - \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot\mathbf{B}(\mathbf{r})\Big]\,\psi \;=\; 0\end{aligned} \tag{6}$$

where the gradient of V and the curl of $\mathbf{A}$ have been written explicitly as electric ($\mathbf{E}$) and magnetic ($\mathbf{B}$) fields. In the non-relativistic limit the first three terms give the ordinary Schrödinger equation. The fourth term gives the relativistic correction due to the change of the mass with velocity (mass-velocity term), while the $\mathbf{E}(\mathbf{r})\cdot\mathbf{p}$ term has no classical analogon and is sometimes called Darwin term. Also these two expressions contain no spin-matrices, therefore they are called scalar-relativistic terms. Note, that the mass-velocity term adds an additional complication to the Schrödinger equation, as it contains explicitly the energy. As an approximation, the non-relativistic energy can be substituted here, or a suitably chosen energy parameter. Scalar-relativistic terms are important for heavy elements, e.g. the different color of silver and gold is caused by a shift of the d-band due to scalar-relativistic effects.
In the second line of eq.(6), the last two terms describe the interaction of the spin of the electron with magnetic fields. The last term obviously gives the coupling to an external magnetic field. Introducing the Bohr magneton, $\mu_\mathrm{B} = \frac{e\hbar}{2mc}$ as the unit of magnetic moment, this term can be written as a Zeeman energy $\mu_\mathrm{B}\boldsymbol{\sigma}\cdot\mathbf{B}$. Let us finally consider the term containing the vector product in eq.(6). If we rewrite it as

$$-\frac{\mu_\mathrm{B}}{2mc}\boldsymbol{\sigma}\cdot(\mathbf{E}(\mathbf{r})\times\mathbf{p}) = -\frac{\mu_\mathrm{B}}{2}\boldsymbol{\sigma}\cdot(\frac{1}{c}\mathbf{E}(\mathbf{r})\times\mathbf{v}) = -\frac{\mu_\mathrm{B}}{2}\boldsymbol{\sigma}\cdot\mathbf{B}_0 \tag{7}$$

we see that, apart from a factor $\frac{1}{2}$, this term has also the form of a Zeeman term.[2] This is the interaction of the electron's spin with the magnetic field that arises in its local frame due to its own motion through an electric field (e.g. from the nuclear potential). In an atom this term couples the electrons spin to its orbital *motion* and is, therefore, called the spin-orbit interaction term.

[2]This factor $\frac{1}{2}$ is called Thomas factor [2], it is frequently discussed for the orbital motion of the electron around the nucleus, see e.g. ref. [3]. However, we have to keep in mind that we are mainly interested in the motion of the electron through a crystal, i.e. also in linearly moving electrons, where this factor is also present.

n	l	s	j	$E_{\mathrm{NR}}[\frac{Z^2}{2}]$	$E_{\mathrm{SOC}}[\frac{\alpha^2 Z^4}{2}]$
1	0	1/2	1/2	1	1/4
2	0	1/2	1/2	1/4	5/64
2	1	-1/2	1/2	1/4	5/64
2	1	1/2	3/2	1/4	1/64
3	0	1/2	1/2	1/9	9/324
3	1	-1/2	1/2	1/9	9/324
3	1	1/2	3/2	1/9	3/324
3	2	-1/2	3/2	1/9	3/324
3	2	1/2	5/2	1/9	1/324

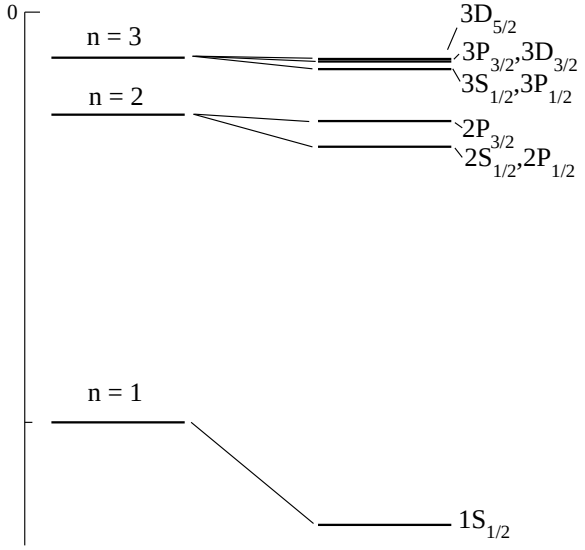


Fig. 2: *Left: table of the non-relativistic (NR) and spin-orbit coupling (SOC) contribution to the energy of an electron in a central potential. Note, that the energy depends just on n and j. Right: illustration of the splittings without SOC contribution (only dependent on n) and with SOC effects (right) in the limit of large Z. The total angular momentum is included in the term symbol as a subscript, e.g. $3P_{\frac{1}{2}}$ means that $n = 3$, $l = 1$ and $j = \frac{1}{2}$.*

2.2 The hydrogenic atom

To investigate the term discussed in eq. (7) in more detail, it is useful to study its effect for a single electron in a central potential, the so-called hydrogenic atom. If we assume that the electric field is derived from a spherically symmetric potential, $V(r)$, (as occurs in the vicinity of an atomic nucleus) we can transform this expression

$$\boldsymbol{\sigma} \cdot (\mathbf{E}(\mathbf{r}) \times \mathbf{p}) = \boldsymbol{\sigma} \cdot (\nabla V(r) \times \mathbf{p}) = \frac{1}{r}\frac{dV(r)}{dr}\boldsymbol{\sigma} \cdot (\mathbf{r} \times \mathbf{p}) = \frac{1}{r}\frac{dV(r)}{dr}(\boldsymbol{\sigma} \cdot \mathcal{L}) = \xi \boldsymbol{\sigma} \cdot \mathcal{L}, \quad (8)$$

with the spin-orbit coupling constant ξ. Note here, that the magnetic field, $\mathbf{B} = \frac{1}{c}(\mathbf{E} \times \mathbf{v})$, is parallel to the orbital momentum, $\mathcal{L}$. Therefore, the spin- and orbital momentum are coupled antiparallel (fig. 1), a fact that is sometimes called third Hund's rule. However, that in contrast to the first and second Hund's rules, which are empirical observations [4] that can be motivated by quantum-mechanical arguments, but not generally proved, this rule is a direct consequence of relativistic quantum-mechanics, at least in the single particle case.

Since the radial derivative of the potential in a crystal will be largest in the vicinity of a nucleus, we can expect that the major contribution to the spin-orbit interaction will come from this region. For small r the potential in an atom will be Coulomb-like ($V(r) = -\frac{Z}{r}$), its derivative $\frac{\partial V}{\partial r}$ is proportional to the nuclear number of the atom, Z. We thus expect that ξ will be large for heavy atoms, but small for lighter ones. Since also the shape of the wavefunction depends on Z, spin-orbit coupling effects that are proportional to $\bar{\xi} = \langle\psi|\xi|\psi\rangle$ grow roughly with the square of Z. E.g. we find for two atoms in the same column of the periodic table, W ($Z = 74$) and Mo ($Z = 42$) that $\bar{\xi}_{\mathrm{W}}/\bar{\xi}_{\mathrm{Mo}} = 3.6$ (data from Ref. [5]).

Suppose we have only one electron characterized (without SOC) by quantum numbers n, l and s, the latter two as given by the eigenvalues of the operators of the orbital angular momentum, $\mathcal{L}$, and the spin $\mathcal{S} = \frac{1}{2}\boldsymbol{\sigma}$. The operator of the total angular momentum is $\mathcal{J} = \mathcal{L} + \mathcal{S}$. It can be

shown that $\mathcal{J}^2$ commutes with a Hamiltonian including a spin-orbit coupling term of form (8) since

$$\boldsymbol{\sigma} \cdot \mathcal{L} = 2\mathcal{S} \cdot \mathcal{L} = \mathcal{J}^2 - \mathcal{L}^2 - \mathcal{S}^2 \tag{9}$$

and that the eigenvalue of $\mathcal{J}$ is $j(j+1)$. Thus j is a good quantum number for an electron in a central potential including spin-orbit coupling and for an hydrogenic atom in the Pauli theory the eigenvalues are given by [1]

$$E = E_{\mathrm{NR}} + E_{\mathrm{SOC}} = -\frac{1}{2}\frac{Z^2}{n^2} - \frac{\alpha^2 Z^4}{2n^3}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right) \tag{10}$$

where n is the principal quantum number and α is the finestructure constant. We further split the energy in a non-relativistic (NR) and spin-orbit coupling (SOC) part. For a single electron we find that j can assume the values $l \pm \frac{1}{2}$ (for $l > 0$, otherwise $j = \frac{1}{2}$), so that the energy spectrum is determined by n, l, s and j as illustrated in figure 2. We see that the total energy depends just on n and j. From this analysis we also find that the energy contribution from SOC scales with the fourth power of the nuclear number, Z, indicating that spin-orbit coupling effects will be more important for heavy atoms.

Let us now consider the case of an atom in a magnetic field, $\mathbf{B}$: In the presence of a magnetic field the levels split according to spin (m_s) and orbital momentum (m_l) and j is no longer a good quantum number. Therefore, we are left with l, s and $m_j = m_l + m_s$ to characterize the atom. From the splitting that results from the interaction of the magnetic moment with an applied field $\mathbf{B}$ it is possible to characterize the electronic state of an atom experimentally. Assume we probe a $2\mathrm{P}_{\frac{3}{2}}$ state: possible values for m_j are $\frac{3}{2}, \frac{1}{2}, -\frac{1}{2}$ and $-\frac{3}{2}$. Since the interaction energy with the $\mathbf{B}$ field is $E_B = \mu_{\mathrm{B}} m_j B$, the state will split in four equally spaced levels (Zeeman effect). This situation will hold, if the magnetic field is weak compared to the coupling between the spin and the orbital moment (fig. 3, left). If, on the other hand, the magnetic field is much stronger than the coupling between spin- and orbital moment, m_s and m_l will couple individually to the $\mathbf{B}$ field and the interaction energy is given by $E_B = \mu_{\mathrm{B}}(m_l + 2m_s)B$. In the $2\mathrm{P}_{\frac{3}{2}}$ state, where available (m_l, m_s) combinations are $(1, \frac{1}{2}), (0, \frac{1}{2}), (0, -\frac{1}{2})$ and $(-1, -\frac{1}{2})$ this leads to a level splitting with energy separations different from the Zeeman effect. In the considered example the separations show a ratio $1 : 2 : 1$ and this strong-field case is called Paschen-Beck effect (fig. 3, right)

If there are more than one electrons involved, in absence of relativistic effects, the electronic structure of an atom is described by the Schrödinger equation with a radially symmetric potential:

$$\sum_i \left\{ -\frac{1}{2}\boldsymbol{\nabla}_i^2 - \frac{Z}{r_i} + \frac{1}{2}\sum_{j \neq i}\frac{1}{r_{ij}} \right\} \Psi(\mathbf{r}_1, \mathbf{r}_2, \dots) = E\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots) \tag{11}$$

where r_i is the distance of electron i from the nucleus and r_{ij} is the mutual distance between electron i and j. For simplicity we use here again atomic units where the mass and the charge of the electron are unity and $\hbar = 1$. Since all parts of the Hamiltonian are rotationally invariant, the Hamiltonian commutes with the operator of the orbital momentum, $\mathcal{L}$, and one of it's components, e.g. $\mathcal{L}_z$. Therefore, we can characterize all eigenfunctions of the Hamiltonian (11) by quantum numbers L and M_L, irrespective of whether we have a single electron or a multi-electron wavefunction. In addition we have, in a nonrelativistic theory, quantum numbers for the total spin, S, and it's projection on an axis, M_S.

In the limit of small spin orbit coupling it is possible to combine first the spin- and orbital momenta to $S = \sum_i s_i$ and $L = \sum_i l_i$, and then work with J, L and S as outlined above

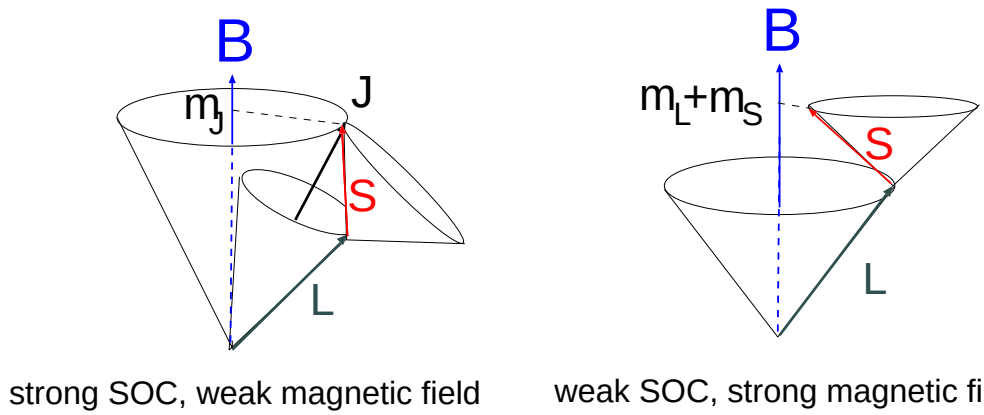


Fig. 3: *Schematic illustration of a* $\mathbf{B}$ *field interaction with an electron of orbital momenta* L, S *and* J*: if the field is weak,* L *and* S *couple to a total moment* J *that precesses around the magnetic field (Zeeman effect). In the Paschen-Beck effect (right) the magnetic field is so strong that* L *and* S *couple individually to* $\mathbf{B}$ *giving rise to a different splitting in the field.*

(Russell-Saunders or ls-coupling scheme). In the case of strong spin orbit coupling the ls-coupling scheme is no longer valid and the spin- and orbital momenta will first couple to a angular momentum j, then the j's will combine to form a total momentum J for the multi-electron system (jj-coupling).
When spin orbit coupling is included, we find that M_S and M_L are no longer constants of motion. From the values of S and L we can deduce the spin-, orbital- and and total magnetic moment

$$\boldsymbol{\mu}_{\mathrm{orb}} = -\mu_{\mathrm{B}}\mathbf{L} \quad ; \qquad \boldsymbol{\mu}_{\mathrm{spin}} = -2\mu_{\mathrm{B}}\mathbf{S} \quad ; \qquad \boldsymbol{\mu}_{\mathrm{tot}} = \boldsymbol{\mu}_{\mathrm{orb}} + \boldsymbol{\mu}_{\mathrm{spin}} = -\mu_{\mathrm{B}}(\mathbf{L} + 2\mathbf{S})\,. \quad (12)$$

From this, it is clear that the total moment is not collinear to $\mathbf{J}$ and can be thought as precessing around $\mathbf{J}$. The component parallel to $\mathbf{J}$ is then [6]

$$|\boldsymbol{\mu}_{\mathrm{tot}\|}| = g_J\mu_{\mathrm{B}}\sqrt{J(J+1)} \quad ; \qquad g_J = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \quad (13)$$

where g_J is called the Landé splitting factor. For the z-component of the total moment we get

$$(\boldsymbol{\mu}_{\mathrm{tot}\|})_z = -g_J\mu_{\mathrm{B}}M_J\,. \quad (14)$$

Again, the quantum numbers L, S and M_J characterizing a particular state can be seen from the magnetic field dependence of the splitting in the Zeeman or Paschen-Beck effect as outlined above.

3 Spin-orbit effects in non-magnetic solids

In contrast to the atomic case, the valence electrons in a solid will arrange to optimize the chemical bonding, e.g. in a simple cubic lattice p_x, p_y and p_z states will form. The level splitting is then determined by the crystal field. Partially, spin-orbit coupling will interfere and lead to additional level splittings as can be observed e.g. in semiconductors at the center of the Brillouin-zone (Γ-point): In Ge without SOC included in the calculation (cf. figure 4, left), we

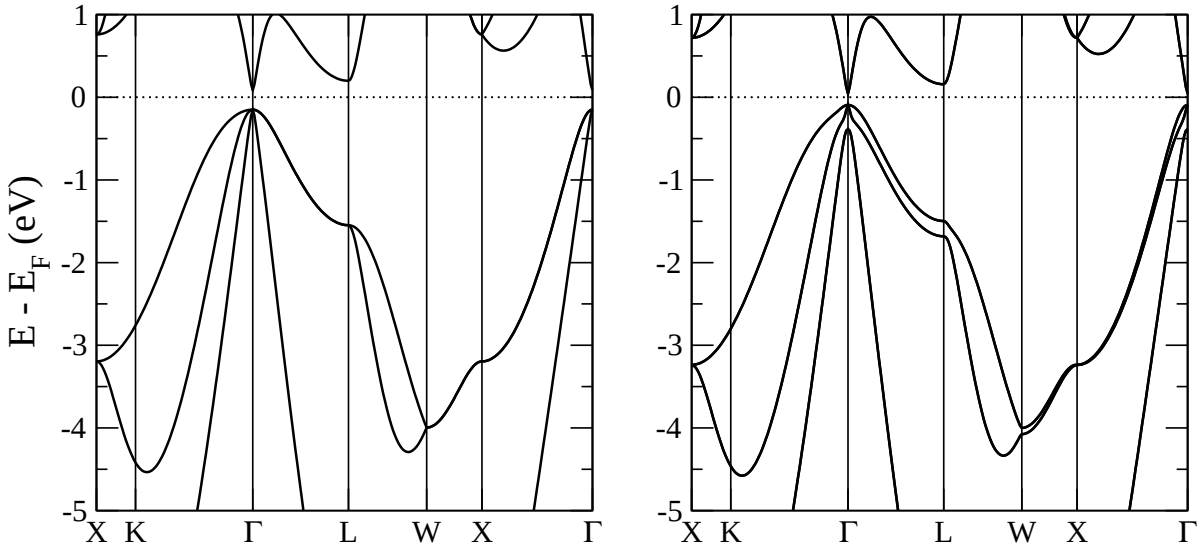


Fig. 4: *Bandstructure of Ge around the Fermi level without spin-orbit coupling (left) and with spin-orbit coupling included (right). Notice, that the three-fold degeneracy of the highest occupied state at the* Γ *point is split by spin-orbit coupling, as well as the doubly degenerate band along the lines* $\overline{\Gamma L}$ *and* $\overline{\Gamma X}$*. The calculation is performed at the experimental lattice constant using the generalized gradient approximation to density functional theory (DFT). Note, that the experimentally observed bandgap of* 0.75 *eV almost closes in a DFT calculation.*

see a three-fold degenerate state directly below the Fermi-level that splits, once SOC is included, into a doubly degenerate and a singly degenerate one. The former one, closer to the Fermi level, in turn consists of two bands with different dispersions, the highly dispersive state is called the light-hole band, the other one is termed heavy hole band. The singly degenerate state at Γ forms the spin-orbit split-off band. In a non-relativistic calculation these bands are degenerate in some high symmetry directions, but when spin-orbit coupling is included a splitting can be observed. As expected, this splitting is smaller in Si, but larger in the isoelectronic but heavier Ge (figure 4, right).

While at the Γ-point, where the crystal momentum is zero, the effect of SOC in solids resembles qualitatively the situation in atoms (quantitatively the effects are smaller, as discussed above), away from the Γ-point new SOC effects arise from the momentum of the electron. These depend strongly on the symmetry of the crystal and we will analyze the "easy" case in detail, while for some cases only the results will be indicated.

3.1 The Rashba- and Dresselhaus effect

In a system without internal or external magnetic field time-reversal symmetry holds, i.e. changing the direction of the arrow of time will not alter the properties of the system. The transformation $t \to -t$ exchanges a particle moving with momentum $\mathbf{k}$ with a particle moving in $-\mathbf{k}$. Time reversal will also invert the precessional motion of the electron and, therefore, its spin. As a consequence, the energy of a right-moving spin-up particle will equal the energy of a left moving spin-down particle,

$$\varepsilon(\mathbf{k},\uparrow) = \varepsilon(-\mathbf{k},\downarrow)\,. \tag{15}$$

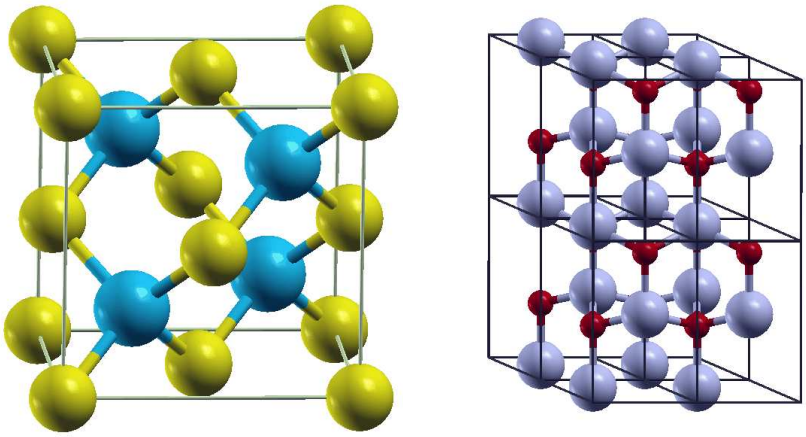


Fig. 5: *Crystal structure of GaAs (zinc blende, left) and ZnO (wurzite, right). Both structures possess no center of inversion, the local symmetry of the atoms in zinc blende structure is nevertheless higher than for the atoms in the wurzite. This leads to different k-dependent spin-orbit splittings, i.e. third-order for the former (Dresselhaus effect) and first order for the latter (Rashba effect).*

In a crystal with inversion symmetry, additionally $\varepsilon(\mathbf{k}) = \varepsilon(-\mathbf{k})$ holds, both for spin-up and spin-down electrons. This means, that the bandstructure is symmetric around the center of the Brillouin-zone, $\mathbf{k} = 0$, and all bands are doubly degenerate. E.g. in the bandstructures in figure 4 show this degeneracy.
In contrast, in crystals without inversion symmetry (for example GaAs) the degeneracy of the bands can be lifted as a consequence of spin-orbit coupling and only eq. (15) holds. This can be understood if we realize that a lack of inversion symmetry, $V(\mathbf{r}) \neq V(-\mathbf{r})$, will result in a non-vanishing potential gradient or electric field, $\mathbf{E}(\mathbf{r})$. As we have seen in the last section an electron moving in an electric field will experience this field Lorentz-transformed as $\mathbf{B}$-field and

$$\varepsilon(\mathbf{k}, \uparrow) \neq \varepsilon(\mathbf{k}, \downarrow)\,. \tag{16}$$

This will, depending on symmetry, result in different consequences for the bandstructures.
That spin-orbit coupling may have important consequences for the one-electron energy levels in bulk semiconductors was first emphasized by Dresselhaus *et al.* [7] already in 1955. Unlike the diamond structure of Si and Ge, the zinc blende structure (figure 5, left), in which for example the III-V semiconductor crystallize, exhibit a bulk inversion asymmetry (BIA), i.e. this crystal structure lacks a center of inversion, so that we can have a spin splitting of the electron and hole states at nonzero wave vectors $\mathbf{k}$ even if $\mathbf{B} = 0$. This is called the Dresselhaus effect. In general, the expressions for the BIA splitting depend on the bands and are quite complicated, e.g. for the lowest conduction band (denoted Γ_6^c) we find a corresponding Dresselhaus Hamiltonian [8]

$$\hat{H}_{\mathrm{D}} = \alpha_{\mathrm{D}} \left[\sigma_x p_x \left(p_y^2 - p_z^2\right) + \sigma_y p_y \left(p_z^2 - p_x^2\right) + \sigma_z p_z \left(p_x^2 - p_y^2\right)\right] \tag{17}$$

with a Dresselhaus constant α_{D} that is small if lighter elements are present in the semiconductor (e.g. for GaAs it amounts to 27.6 eV/Å^3) while heavier elements show larger values (760.1 eV/Å^3 for InSb). Since this Dresselhaus term produces a spin splitting which is proportional to k^3, $\varepsilon_{\mathrm{D}} \propto k^3$, it is rather small in the relevant region around Γ.

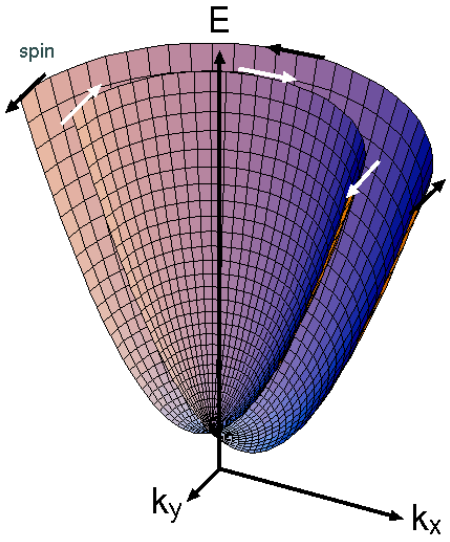


Fig. 6: *Cut through the parabolic energy dispersions of a two-dimensional electron gas in a structure inversion asymmetric (SIA) environment. Indicated are the vector fields of the spin-quantization axes (or the patterns of the spin) at the Fermi surface. As the opposite spins have different energies, the Fermi surface becomes two concentric circles with opposite spins. The effective B-field,* $\mathbf{B}_{\text{eff}}$ *is always perpendicular to the propagation direction defined by* $\mathbf{k}_{\|}$.

While the zinc blende structure thus allows mainly for spin-splitting effects that are third-order in k, in crystals with wurzite structure (figure 5, right) also terms linear in k are allowed. The symmetry of these crystals was analyzed by Rashba in 1959 [9], the paper focusing on the analysis including spin-orbit effects was unfortunately never translated [10]. Nevertheless, we can see from the wurzite lattice that it consists of bi-layers different atoms that locally create electric fields in z-direction, leading to the spin splitting linear in k, $\varepsilon_{\mathrm{R}} \propto k$, analogously to the situation on surfaces. Therefore, we will focus on this term in the next section. Recently, this "bulk Rashba-effect" created a lot of interest again, leading to very large spin splittings in compounds of BiTeI type [11].

3.2 Surfaces: the Rashba-Bychkov effect

Performing a Taylor expansion of the potential $V(\mathbf{r})$, $V(\mathbf{r}) = V_0 + e\mathbf{E}(\mathbf{r}) \cdot \mathbf{r} + \cdots$, in lowest order the inversion asymmetry of the potential $V(\mathbf{r})$ is characterized by an electric field $\mathbf{E}(\mathbf{r})$. When electrons with an effective mass m^* propagate with a velocity $\mathbf{v} = \mathrm{d}\varepsilon/\mathrm{d}\mathbf{p} = \frac{1}{m^*}\mathbf{k}$ in an external electric field $\mathbf{E}$ defined in a global frame of reference, then the relativistic Lorentz transformation gives rise to magnetic field $\mathbf{B} = \frac{1}{c}(\mathbf{v} \times \mathbf{E}) = \frac{1}{m^* c}(\mathbf{k} \times \mathbf{E})$ in local frame of the moving electron. The interaction of the spin with this $\mathbf{B}$ field leads then to the so-called Rashba or Bychkov-Rashba Hamiltonian [12, 13]

$$\hat{H}_{\mathrm{R}} = \alpha_{\mathrm{R}} \boldsymbol{\sigma} \cdot (\mathbf{p} \times \mathbf{E}) \qquad \text{or} \qquad \hat{H}_{\mathrm{R}} = \alpha_{\mathrm{R}}\, \boldsymbol{\sigma} \cdot (\mathbf{k} \times \mathbf{E}) \qquad \text{or} \qquad \hat{H}_{\mathrm{R}} = \alpha_{\mathrm{R}}(|\mathbf{E}|)\, \boldsymbol{\sigma} \cdot (\mathbf{k} \times \hat{\mathbf{e}}) \quad (18)$$

describing the Rashba spin-orbit coupling as additional contribution to the kinetic energy. $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices, eq.(2). The latter two terms are strictly correct only for plane wave eigenstates as, e.g. for a two-dimensional electron gas (2DEG). An important realization of a 2DEGs are electrons in doped semiconductor heterostuctures, that support an electron gas at the interface between two materials, e.g. (InGa)As and InP [14]. Another possibility to study the Rashba-effect in 2DEGs is provided by angle-resolved photoemission on surfaces which support a surface state, e.g. in Au(111) [15]: The electrons of the surface state move in a potential gradient that is provided by the surface itself (but can also be modified slightly by external electric fields [16]).
The general features of the Rashba-model are studied for the 2DEG in a potential with structural inversion asymmetry (SIA) and the corresponding bandstructure are displayed schematically in

figure 6. For electrons propagating in the 2DEG extended in the (x, y) plane subject to an electric field normal to the 2DEG, $\hat{\mathbf{e}}_z = (0, 0, 1)$, the Hamiltonian takes the form

$$\hat{H} = \hat{H}_{\rm kin} + \hat{H}_{\rm R} = \frac{\mathbf{p}_\parallel^2}{2m^*} + \alpha_{\rm R} \left(\boldsymbol{\sigma} \times \mathbf{p}_\parallel\right)_{|z} = \frac{\mathbf{p}_\parallel^2}{2m^*} + \alpha_{\rm R} \left(\sigma_x p_y - \sigma_y p_x\right), \tag{19}$$

which can be solved analytically. For a Bloch vector in the plane of the 2DEG, $\mathbf{k}_\parallel = (k_x, k_y, 0) = k_\parallel(\cos\varphi, \sin\varphi, 0)$, the eigenstates written as a product of plane wave in space and two-component spinor are

$$\psi_{\pm\mathbf{k}_\parallel}(\mathbf{r}_\parallel) = \frac{e^{i\mathbf{k}_\parallel \cdot \mathbf{r}_\parallel}}{2\pi} \frac{1}{\sqrt{2}} \begin{pmatrix} ie^{-i\varphi/2} \\ \pm e^{i\varphi/2} \end{pmatrix} \tag{20}$$

with eigenenergies

$$\varepsilon_\pm(\mathbf{k}_\parallel) = \frac{\mathbf{k}_\parallel^2}{2m^*} + \alpha_{\rm R} \left(\boldsymbol{\sigma} \times \mathbf{k}_\parallel\right) = \frac{\mathbf{k}_\parallel^2}{2m^*} \pm \alpha_R |\mathbf{k}_\parallel| = \frac{1}{2m^*}(k_\parallel \pm k_{\rm SO})^2 - \Delta_{\rm SO}\,, \tag{21}$$

where $\pm$ denotes the spin-up and -down states with respect to a spin orientation axis $\hat{\mathbf{n}}(\mathbf{k}_\parallel)$, local in $\mathbf{k}_\parallel$ space. With the exception of the high-symmetry state $k_\parallel = 0$, we find that the original two-fold degenerate energy paraboloid of the 2DEG in a constant potential is indeed spin-split. This splitting $\varepsilon_+(\mathbf{k}_\parallel) - \varepsilon_-(\mathbf{k}_\parallel) = 2\alpha_{\rm R} k_\parallel$ is linear in $k_\parallel$. Due to the presence of the SIA potential and the spin-orbit interaction, the origin of the degenerate parabola is shifted by $k_{\rm SO} = m^*\alpha_{\rm R}$, but in opposite directions for up- and down-spins with in overall spin-orbit lowering of $\Delta_{\rm SO} = m^*\alpha_{\rm R}/2$. The orientation axis is given by the expectation value

$$\hat{\mathbf{n}}_\pm(\mathbf{k}_\parallel) = \langle \psi_{\pm\mathbf{k}_\parallel} | \boldsymbol{\sigma} | \psi_{\pm\mathbf{k}_\parallel} | \rangle = \pm \begin{pmatrix} \sin\varphi \\ -\cos\varphi \\ 0 \end{pmatrix} \quad \perp \quad \mathbf{k}_\parallel = k_\parallel \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 0 \end{pmatrix}. \tag{22}$$

We find that the orientation axis is independent of the magnitude $k_\parallel$ and depends only on the direction of the $\mathbf{k}_\parallel$ vector. In fact, it is in the plane of the 2DEG and the orientation axis is perpendicular to the propagation direction of the electron. Considering $\mathbf{k}_\parallel \to -\mathbf{k}_\parallel$, φ changes to $\varphi + \pi$, we find that the spin orientation axis reverses as indicated in figure 6. Thus for $\mathbf{k}_\parallel$ and $-\mathbf{k}_\parallel$ the spin-up and -down states refer to opposite orientations. Defining a global quantization axis along the line $(-\mathbf{k}_\parallel, \mathbf{k}_\parallel)$, e.g. according to $\hat{\mathbf{n}}_\pm(+\mathbf{k}_\parallel)$, then a spin-up state appears as spin-down state if $\mathbf{k}_\parallel$ changes sign. Together with the eigenvalue spectrum given in equation (21) the Kramer degeneracy $\varepsilon_\uparrow(\mathbf{k}_\parallel) = \varepsilon_\downarrow(-\mathbf{k}_\parallel)$ holds. In all, the magnetic moment is zero when averaged over all states $\mathbf{k}_\parallel$. This is consistent with the absence of an $\mathbf{B}$ field.
The Rashba spin-orbit splitting may be observed either by electron photoemission or transport experiments. Transport experiments work typically on a shell of constant energy ε. Here we expect at a given energy two different wave vectors for up- and down-electrons, which will be exploited in the Datta-Das proposal of a spin-transistor [17].

4 Topological Insulators

Let us start with a simple, not very accurate analogy: if we consider a bulk ferroelectric material, e.g. $BaTiO_3$, it is at first sight just an insulator in a crystal structure without inversion symmetry. Conceptually, the ions in the crystal can be grouped to dipoles, but otherwise this

has no direct consequence. However, when we introduce a boundary (e.g. a surface or an interface to a non-ferroelectric material or a ferroelectric with opposite polarization) a charged layer at this boundary appears and the (bulk) polarization, P, determines the amount of charge. If we consider a head-to-head domain wall in a ferroelectric material the plane, where P changes to $-P$, forms a two-dimensional metallic boundary between two insulators.
Of course there only a few ferroelectrics among all insulators. More fundamentally, insulators with time-reversal symmetry (i.e. nonmagnetic ones without external magnetic field) can be characterized by a topological number, ν, that bears consequences that a very similar to our example of the ferroelectric polarization: At the boundary between two materials with different ν a metallic interface has to form. In contrast to the charged layer in ferroelectrics, which is easily destroyed by defects, this metallic layer is robust and protected as long as time-reversal symmetry is not broken.
This topological description of matter had of course a predecessor, most notably the description of the quantum Hall effect (QHE) by Thouless et al. [18]. There, the conductivity at the edge of two-dimensional electron gas in a strong magnetic field is given by the so called Chern number. As we will see below, in a non-magnetic insulator the spin-orbit coupling can take now the role of the external magnetic field and it is the topological number ν that characterizes the state of matter [19].
After its discovery in the year 2005 the field of topological insulators (TIs) experienced an enormous expansion and created unforeseen activity in the solid-state physics community and beyond. Although nowadays the focus is on heavy materials with strong SOC effects, like bismuth and mercury compounds, the field started conceptually on the other end of the periodic table, with graphene. To introduce some basic concepts, like the quantum spin Hall effect (QSHE) we will start our discussion with this material, before we turn to popular 2- and 3-dimensional TIs. A careful comparison between graphene and TIs shows that there are some special properties of graphene that have no analog in other TIs (and vice-versa), therefore graphene is sometimes called a marginal topological insulator [20].

4.1 Graphene - a marginal topological insulator

The simplicity of the electronic structure of graphene, a single layer of carbon forming a honeycomb lattice, makes it attractive to discuss TIs first on this example, although practically the effects are much too small to be detected experimentally. The Fermi-surface of graphene consists just of two points, K and K' (see fig. 7). DiVincenzo and Mele realized, that the electronic structure in the vicinity of these points can be described rather simply [21]: There are two atom types (A, B) in the honeycomb lattice and at the K-point the p_z wavefunctions on these atoms are independent from each other. Therefore, orbitals at A and B can be taken as a basis that be written in spinor form. This so-called pseudo-spin is in the graphene community usually denoted σ, while s is the spin. Here, we have to adapt the opposite convention so that in that basis, the Hamiltonian takes the simple form:

$$\left[\frac{W}{i}\mathbf{s}\cdot\nabla + U(r) - E\right]\Psi(r) = 0 \tag{23}$$

where W is related to the bandwidth and $U(r)$ can be some external potential. The the components of $\mathbf{s}$ are Pauli matrices, here for the pseudo-spin. If we take $U(r)$ as a staggered potential, acting oppositely on sites A and B, we can see that this equation has the same structure as the

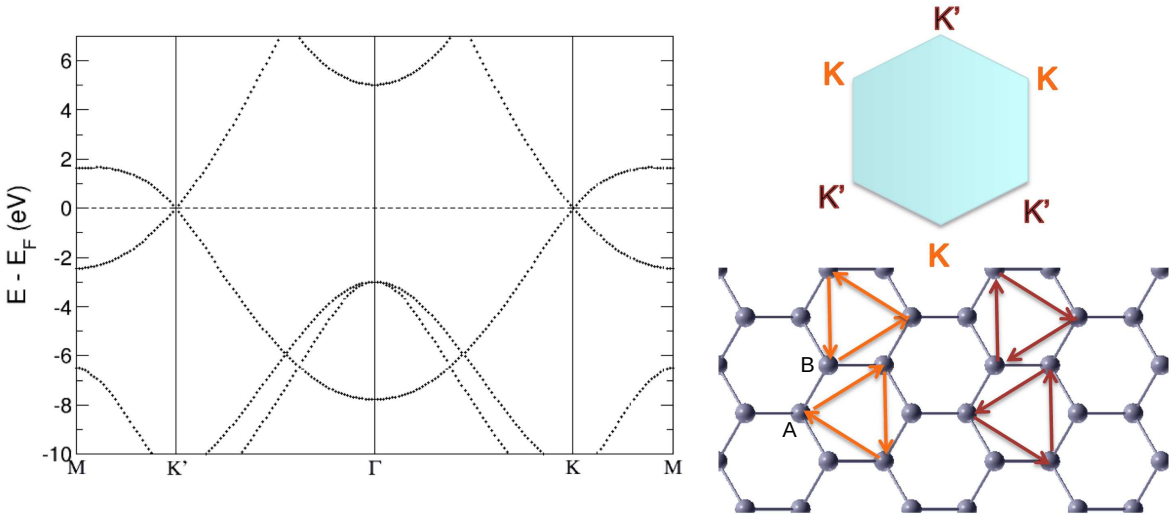


Fig. 7: *Bandstructure of graphene (left) and the Brillouin-zone (upper right) indicating the position of the K and K' points. Graphene has a bipartite lattice with atom types labeled A and B (lower right). Hopping processes for electrons with momentum K and K' are indicated by arrows.*

two-dimensional Dirac equation with the Hamiltonian

$$\hat{H}^{\mathrm{2D}} = c\boldsymbol{\sigma}\cdot\mathbf{p} + mc^2\sigma_z; \qquad \text{where } \boldsymbol{\sigma} = (\sigma_x, \sigma_y) \qquad \text{and } \mathbf{p} = (p_x, p_y)\,. \tag{24}$$

In total there are now eight states available at the Fermi level, E_{F}: each combination of spin, pseudo-spin, and the so-called valley (K or K') characterizes one level. Possible hoppings for electrons in these states are shown on the lower right of fig. 7.
Now let us consider the effect of SOC on these electronic states: We distinguish in fig. 7 two types of hoppings: for an electron that hops among the A-sites with momentum K the nearest B-site is on the left of its path. Similarly, for hopping on the B-site with momentum K' [in short (B, K')] the nearest A-site is on the left. For the other states (A, K') and (B, K) the nearest neighbor sites are on the right. If we think of the influence of the potential gradient of the nearest neighbor atom (electric field $\mathbf{E}$) on an electron moving with momentum $\mathbf{k}$ as a Rashba-field $\propto \mathbf{E}\times\mathbf{k}$, it is clear that SOC will lower the energy for one type of states, (A, K) or (B, K'), for one spin direction (+), for the other states, (A, K') or (B, K), the other spin direction (−) will be preferred. Kane and Mele [22] cast this into the Hamiltonian:

$$\hat{H}_{\mathrm{SOC}} = \Delta_{SO}\tau_z\sigma_z s_z \tag{25}$$

where τ_z is a Pauli matrix that refers to the valley (K, K'), σ to the spin and s to the pseudo-spin, while Δ_{SO} gives the spin-orbit coupling strength. Actually, the SOC splitting of graphene at the K-point is tiny (24μeV, see e.g. [23] or fig. 8, left) and the reason will become clear later.
As next, most subtle point we have to consider the actual hopping process: analysis with a tight-binding model shows, that the direct next-nearest-neighbor hopping does not contribute significantly. There are rather two nearest-neighbor (NN) hops, t_{NN}, involved. As the NN hopping from a p_z to a p_z orbital is forbidden due to the different phases of the orbitals, more complex processes have to be involved. Let us consider a hopping via d-orbitals at the NN site [24]:

$$\left|p_z^{\mathrm{A}}\right\rangle \rightarrow \left|d_{xz}^{\mathrm{B}}\right\rangle \rightarrow \left|d_{yz}^{\mathrm{B}}\right\rangle \rightarrow \left|p_z^{\mathrm{A}'}\right\rangle\,. \tag{26}$$

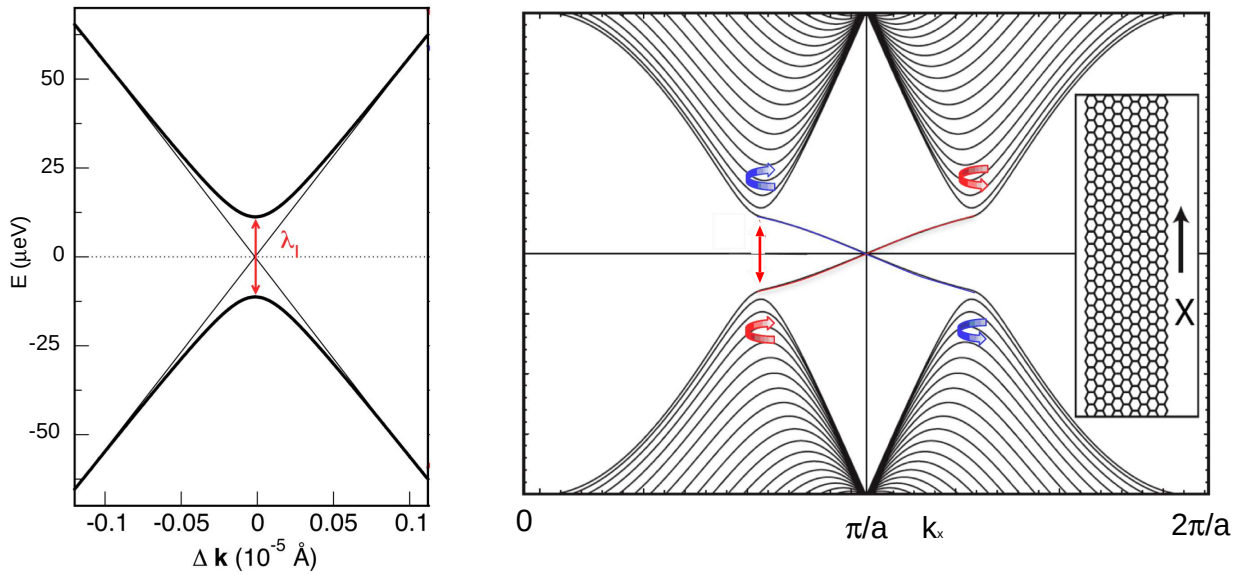


Fig. 8: *Bandstructure of graphene near the K-point calculated in DFT with SOC included (left). The intrinsic spin-orbit splitting λ_{I} is 24μ eV. Model bandstructure of a (H-terminated) graphene ribbon with spin-polarized edge states (red and blue indicate the spin-direction) crossing the gap (adapted from [22]).*

While the first and the last arrow in this transition indicates simple $pd\pi$ NN hopping, for the transition from d_{xz} to d_{yz} we have to involve SOC again. It is a spin-conserving process, but $\langle d_{xz}|H_{\mathrm{SOC}}|d_{yz}\rangle = -i\sigma_z$ introduces a complex phase in the hopping process. This complex phase is actually a key ingredient to the famous Haldane model, proposed to realize a "'Quantum Hall Effect without Landau Levels"' in a solid [25]. It is exactly this phase that leads in his model to a quantized edge conductivity σ^{xy}. For more details refer to the lecture on Berry phases by Y. Mokrousov (A6) and next subsection.
For our proposes it is sufficient to highlight the role of SOC, causing a kind of intrinsic rotation of the electron via the hopping process. The sense of rotation is of particular importance: all electrons of one spin direction pick up one sense of rotation, all electrons of the other spin direction rotate in the opposite way. This results in a total orbital current, that is opposite for the two spin channels. In the bulk of the material, this current is (like the ferroelectric polarization) hard to grasp, but it can be cast into a topological number (a spin Chern number). On the boundary, however, it results in quantized *spin* conductivity, the Quantum Spin Hall Effect (QSHE). The states that bridges the tiny spin-orbit gap of graphene that carry this spin-current are shown in fig. 8, right.
Many basic features of TIs can already be derived from the bandstructure of the edge-states of graphene:

- The bands cross a spin-orbit gap, that is usually small (despite the name "insulator").
- They are spin-polarized bands, the splitting is basically Rashba-like, but
- for a given energy there is only one band per spin available.

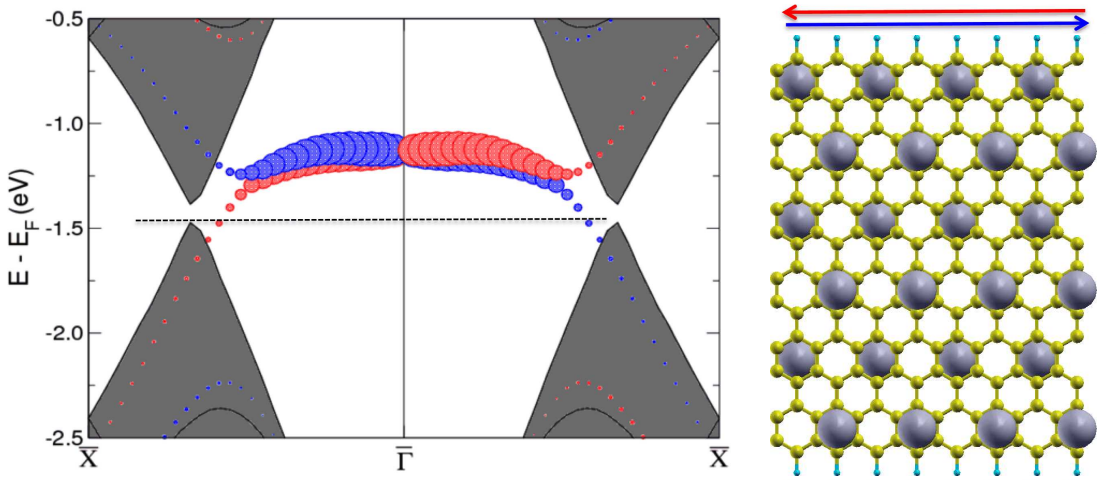


Fig. 9: *Bandstructure of a H-terminated graphene nanoribbon with Tl adatoms calculated in DFT with SOC included (left). The size of the red and blue symbols indicates the localization of electrons of opposite spin-direction at the edge of the ribbon, leading to edge transport. The Tl atoms are distributed equally above and below the ribbon as indicated in the right figure.*

The latter fact leads to the robustness of these edge-states, in first approximation we can say that the scattering from k to $-k$ is prohibited since this would involve a reversal of the spin of the electron. More strictly, it can be shown that scattering is prohibited between time-reversal partner $(k, \uparrow)$ and $(-k, \downarrow)$ [26]. In contrast to other TIs, the edge state in graphene is also present without SOC and was already discovered for H-terminated graphene ribbons by Fujita et al. in tight-binding calculations [27].

We should note that many of the features listed above, that characterize a TI, can also be present in a Rashba-type SOC system. E.g. in an external magnetic field perpendicular to the surface, the band crossing at the origin ($k_x = k_y = 0$ in fig. 6) opens and if the Fermi level falls into that gap, only one band per spin is available for transport. This is a scenario often discussed in the context of coupling to superconductors for the realization of states called Majorana fermions [28]. Nevertheless, these states are not robust against perturbations at the edge of the sample, in that sense they differ from the characteristics of a TI. Therefore, one has to be careful: absence of backscattering, spin-momentum locking etc. are no sufficient proof for topological behavior in an investigated system.

Before we continue the discussion of two- and three-dimensional TIs, let us shortly comment on the possibility to enhance SOC effects in graphene. In principle, the vicinity of heavy elements (e.g. when they are used as substrate) is known to enhance SOC effects in graphene. Graphene deposited on Au(111) shows a splitting of the p_z bands of 10 meV [29], additional adatoms can enhance this value to about 100 meV [30]. This value was measured using spin- and angle-resolved photoemission spectroscopy (SP-ARPES) slightly off the K-point in the experimentally accessible energy range. On the other hand, on a light substrate, like SiC, only about 50 μeV splitting are found [31]. However, the induced SOC effects from a substrate are of Rashba-type, i.e. the spin-orbit field is in-plane, while the intrinsic SOC induces a spin-polarization of the bands that is oriented perpendicular to the graphene plane. This interplay of intrinsic and Rashba-type SOC does not split the Dirac cone if the latter interaction is larger than the former, leading to a complex pattern of spin-polarized bands [23]. Only a careful balance

of SOC-inducing neighboring elements can restore the right symmetry for a QSHE in graphene with larger bandgap (see fig. 9 or [32]).
Finally, we want to point out the similarity of graphene and the bismuth (111) bilayer: the latter one can be seen as a corrugated honeycomb lattice, like graphene but with atoms A and B displaced perpendicular to the layer. The SOC effects in Bi (nuclear number 83) are much larger than in carbon, so we can expect a sizable effect here. Indeed, the Bi bilayer was proposed as a TI already one year after the discovery of the QSHE in graphene [33]. But it turns out that these bilayers are structurally not stable and transform into a black-phosphorous structure that is topologically trivial [34]. However, recently progress has been made to stabilize these hexagonal layers by epitaxial growth on the closed-packed surface of Bi_2Te_3, another TI that will be discussed below [35].
On the other hand, there are also significant differences between the Bi (111) bilayer and the graphene sheet: (i) The flatness of graphene restricts the possible hopping mechanisms for the p-orbitals, e.g. between p_z and the $p_{x,y}$ orbitals involved in bonding. The corrugation of the Bi bilayer allows all p-orbitals to contribute to hopping and electronic transport. This additional flexibility is not directly visible in the QSHE, but in the presence of a magnetic (exchange) field, where the quantum anomalous Hall effect (QAHE) can appear, it offers a larger variety of quantized conductive states [36]. (ii) While graphene is a semimetall (i.e. although the density of states vanishes at the Fermi level, the size of the bandgap is also zero) and shows a "mass-inversion" between the K and the K'-point, the Bi bilayer is a "real" insulator (0.2 eV band gap) and the band order is inverted with respect to a normal (trivial) insulator. (We will discuss these different inversions of the mass in the next subsection.) E.g. a (111) bilayer of Sb, which is chemically equivalent but has a smaller spin-orbit coupling strength, is topologically trivial and at the interface between such Sb and Bi bilayers topologically protected edge-states appear. While the Sb bilayer is the topologically trivial equivalent of the Bi bilayer, there is no "trivial equivalent" to graphene.

4.2 Band inversion and topology in 2D topological insulators

Shortly after the prediction of the QSHE in graphene another realization of this effect was proposed where SOC effects (and, hereby, the energy scales) are considerably larger: HgTe/CdTe quantum wells [37]. The choice of this particular material system was motivated by experiments, therefore it was also the first system where the QSHE was demonstrated by measurements of the edge conductance [38].
To show the similarities with the physics discussed in the last chapter, let us briefly recall the graphene Hamiltonian with SOC in the following form:

$$\hat{H} = \nu_{\mathrm{F}} \left(s_x \tau_z p_x + s_y p_y \right) + \Delta_{\mathrm{SO}} \tau_z \sigma_z s_z \,. \tag{27}$$

In the language of the 2D Dirac equation (24) we can say that the spin-orbit term gives a mass to our particles. For a given spin and valley the Hamiltonian can be written:

$$\hat{H}_K = \begin{pmatrix} +\Delta_{\mathrm{SO}} & \nu_{\mathrm{F}} \left(p_x - i p_y \right) \\ \nu_{\mathrm{F}} \left(p_x + i p_y \right) & -\Delta_{\mathrm{SO}} \end{pmatrix} ; \quad \hat{H}_{K'}(k) = \begin{pmatrix} -\Delta_{\mathrm{SO}} & -\nu_{\mathrm{F}} \left(p_x + i p_y \right) \\ -\nu_{\mathrm{F}} \left(p_x - i p_y \right) & +\Delta_{\mathrm{SO}} \end{pmatrix} \tag{28}$$

We see that, going from K to K', the mass-term changes sign, causing an *inversion of the gap*. For the edge state this means that one spin channel connects the valence band at K with the

conduction band at K', the other spin channel crosses from the conduction to the valence band (fig. 8).

The Hamiltonian (28) describes two spin-degenerate bands near the Fermi level separated by a gap. If we compare to the situation in semiconductors like Ge (fig. 4), we see that it is slightly more complicated there: apart from the already discussed heavy- and light hole bands and the split-off band in the valence states, there is a steeply dispersing empty state in the conduction band. There is a long tradition to describe these bands (sometimes labeled HH, LH, SO, and E, respectively) in $k \cdot p$ perturbation theory, e.g. the eight-band Kane model [39] or the Luttinger-Kohn Model [40].

Here, we will look at a much simpler Hamiltonian, taking only the heavy-hole (HH) and empty state (E) into account. In II-VI semiconductors, like CdTe, the E state is mainly derived from the group II (Cd) s-level (so called Γ_6 state), the HH state comes from the group IV (Te) p states that are spin-orbit split into the doubly degenerate Γ_8 and singly degenerate Γ_7 level. In CdTe the Γ_6 state is above the Γ_8, the bandgap is about 1.6 eV [41]. In HgTe, the larger nuclear number of Hg as compared to Cd mainly affects the Γ_6 state: although the Darwin term increases the energy of that level in the heavier element, the mass-velocity term (eq. (6)) more than compensates this increase and finally the energy is 0.3 eV lower than the Γ_8 level [42]. This is a *band-inversion* driven by relativistic effects, notably not SOC in this case. DFT calculations in scalar-relativistic approximation already show this band inversion and even overestimate it due to the underestimation of the band gap in DFT [41].

When we describe the two bands, Γ_6 and the "heavy" part of Γ_8 now in the basis $(\left|\mathrm{E}, m_j = \frac{1}{2}\right\rangle, \left|\mathrm{HH}, m_j = \frac{3}{2}\right\rangle, \left|\mathrm{E}, m_j = -\frac{1}{2}\right\rangle, \left|\mathrm{HH}, m_j = -\frac{3}{2}\right\rangle)$ the effective Hamiltonian reads [37]:

$$\hat{H}_{\text{eff}}(k_x, k_y) = \begin{pmatrix} \hat{H}(\mathbf{k}) & 0 \\ 0 & \hat{H}^*(-\mathbf{k}) \end{pmatrix} \quad \text{with} \quad \hat{H}(\mathbf{k}) = \varepsilon(\mathbf{k}) I_2 + \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma} \tag{29}$$

$$\varepsilon(\mathbf{k}) = C - D(k_x^2 + k_y^2) \quad \text{and} \quad \mathbf{d}(\mathbf{k}) = \begin{pmatrix} Ak_x \\ Ak_y \\ M - B(k_x^2 + k_y^2) \end{pmatrix} \tag{30}$$

where A, B, C, D, and M are material-dependent parameters. Obviously, C defines the valence band offset and M relates to the gap-size. The parameter A is equivalent to the velocity, ν_{F}, in graphene, compare eq. (27). Finally B and D define the quadratic terms of the bands. How these parameters relate to the six-band Kane model is described in detail in the supplement of Ref. [37].

From our description of HgTe and CdTe above, it is clear that the major difference between the Hamiltonians describing these two systems is the sign of M: HgTe has an inverted gap, CdTe has not. Consider now the interface between these two materials. In principle one could simply try to calculate the electronic structure by some method, however topology (cf. contribution A6) offers a more elegant solution.

Many properties of a Hamiltonian can be derived from its topological properties, in our case it is the topology in momentum space [20]. When the Hamiltonian can be written as $\hat{H} = \varepsilon(\mathbf{k}) + \mathbf{d}(\mathbf{k}) \cdot \boldsymbol{\sigma}$, a topological number c can be defined

$$c = \frac{1}{4\pi} \int d^2k \hat{\mathbf{d}} \cdot \left(\frac{\partial \hat{\mathbf{d}}}{\partial k_x} \times \frac{\partial \hat{\mathbf{d}}}{\partial k_x} \right) \tag{31}$$

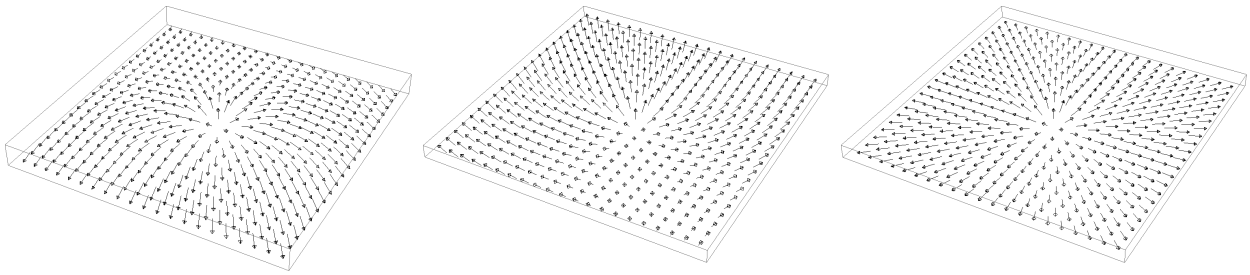


Fig. 10: *Illustration of the unit vector field* $\hat{\mathbf{d}}(\mathbf{k})$ *for* $A = B = 1$ *and* $M = 0.1$ *(left) that gives a topologically non-trivial solution (c=1), while the trivial solution is obtained for* $A = -B = 1$ *and* $M = 0.1$ *(middle). While* $\hat{\mathbf{d}}$ *in the left picture maps on the full unit sphere (skyrmion), on the right only half the unit sphere is covered (meron). In the plots* k_x *and* k_y *range from* -1 *to* 1.

with $\hat{\mathbf{d}} = \mathbf{d}/|\mathbf{d}|$. It defines the mapping from reciprocal space k_x, k_y to the unit vector field $\hat{\mathbf{d}}$. This number can be seen as a Chern number, in real space (magnetic systems) it is sometimes called winding number. A few examples are shown in fig. 10: In case the M and B term have the same sign (left figure), $\hat{\mathbf{d}}$ changes from $+z$ direction for $k_x = k_y = 0$ to $-z$ at large k values. The vector field has the topology of a sphere, $c = 1$, the structure is sometimes called a skyrmion. In contrast, if M changes sign, the vector has the same direction at the origin as at the boundaries. In between it tilts a bit, but topologically this does not matter: since only a part of the unit sphere is covered, the topological number is zero. Finally, we show the case for one K point of graphene (right): $\hat{\mathbf{d}}$ covers half the unit sphere, the other half is covered at the K' point. The resulting structures are sometimes called merons (cf. the contribution on complex magnetism, C4).

Of course our Hamiltonian, eq. (29), contains two spin sectors $\hat{H}_S(\mathbf{k})$ and $\hat{H}_{S'}(\mathbf{k}) = \hat{H}^*_S(-\mathbf{k})$ and the topological numbers of the two sectors have to be opposite. In total, the topological number $c = c_S + c_{S'} = 0$ as required for a time-reversal invariant system (see eq. 29). But we can define a quantity $c_s = \frac{1}{2}(c_S - c_{S'})$ that casts the differences of the band-topology of a system with and without band inversion into a single number.

Let us note one peculiarity of winding numbers characterizing topological insulators, i.e. topological structures induced by spin-orbit coupling effects in time-reversal invariant systems: This is best illustrated showing the contrast to the QHE, induced by an external magnetic field, that can be characterized by a Chern number, which is directly proportional to the number of conducting channels at the edge of the sample. As a function of the magnetic field, this number increases in uniform steps, the Chern number is an arbitrary integer (a $\mathbb{Z}$ number). The current proportional to this number is a charge current. On the other hand, in the QSHE, the edge current is a spin-current that results from oppositely propagating quantized currents for one and the other spin channel. The topological index characterizing this state is either zero or one, i.e. $\nu = 0$ signifies a trivial insulator, $\nu = 1$ a TI. Therefore, this number is called the $\mathbb{Z}_2$ index.

The elegance of this description comes now, if we want to describe the interface between two systems with different topological numbers. Of course, we cannot predict the electronic structure at this interface in detail, but we know that the two topologically different bandstructures cannot be simply connected, very much in the same way as the edge state in graphene cannot simply connect the valence band at K with the valence band at K', but has to cross the gap to at-

tach to the conduction band at K'. In analogy, the two insulators HgTe and CdTe cannot have an insulating interface, some electronic state has to appear in the gap to connect the topologically inequivalent bandstructures. Since the existence of this state is a bulk property, it is not affected by the details of the interface, it is *topologically protected*. Experimentally, the existence of these edge channels was detected in HgTe/CdTe quantum wells, the quantized conductance is a clear signature of the topological origin of these states, like the quantum Hall effect (QHE) is characterized by its quantized Hall plateaus [43]. One has to keep in mind that HgTe itself is a semimetal, i.e. the valence and conduction bands touch at Γ, but in the confined geometry of the quantum well a gap can be opened due to finite-size quantization of the system.
It would lead too far to discuss the fascinating physics that can be described by topological concepts, next to the QSHE and QHE it can be applied to types of superconductivity, ^{3}He, and more examples can be found in the book of Volovik [44]. Here, we restrict ourselves to pointing out the role of relativistic effects that lead to band inversion and a change of the mass term in the Hamiltonian (29). Clearly, there are also other mechanisms that can lead to a change in topology of the electronic structure, but for topological insulators the relativistic effects and, in particular, spin-orbit coupling are most relevant.

4.3 Examples of 3D topological insulators

Up to now we focused on 2D systems: historically, these were the materials where TIs were discovered and their mathematical description is simpler than the corresponding 3D systems. The first 3D TI, that was discovered, was the Bi_xSb_{1-x} alloy. In the semimetals Bi and Sb for all $\mathbf{k}$ points of the Brillouin zone there is a gap in the vicinity of the Fermi level but at E_F small electron and hole pockets form a non-vanishing density of states. These materials are semimetals but since the early seventies, it is known that alloying Sb with Bi produces a semimetal to semiconductor transition and opens a bandgap of about 20 meV. While this value might seem small, for low temperature conductivity experiments it forms no obstacle. We confine our discussion here to pure Bi and Sb, since it is sufficient to determine the topological character of the elements: If alloying transforms the bandstructure of one of these elements from a semimetal to a semiconductor without producing any band crossings, the character of the parent compound is conserved and we know whether the alloy is a TI or not.
Strong Rashba-type SOC effects have been observed in Bi and the spin-polarization of the surface states (on the natural, non-magnetic surface) was confirmed by calculations and experimentally not only on surfaces of bulk material, but also on thin bismuth films [45]. Therefore, it was rather natural to expect the existence of this effect also in the chemically similar antimony. Although the bandstructures of these elements are very similar, there is a subtle difference between the electronic states in Sb and Bi, which only becomes apparent when the symmetry of the wavefunctions is taken into account: At a certain point (L) in the bulk Brillouin zone, the symmetry of the states that form the upper and lower edge of the gap in Sb is reversed as compared to Bi. This exchange of bands is a consequence of the increased spin-orbit coupling in Bi, as can be seen from calculations, where the SOC strength is artificially decreased from the Bi to the Sb value (fig. 11). This little detail in the bandstructure turns out to have severe consequences for the surface states in both materials.
Combining density functional theory calculations with spin- and angle-resolved photoemission experiments it was now possible to pin down the topological nature of Sb and the gaped BiSb alloy [46]. DFT calculations confirm the non-trivial topological ordering of the bands in Sb and the difference in the surface states can also be clearly seen in fig. 11: On the Sb(111)

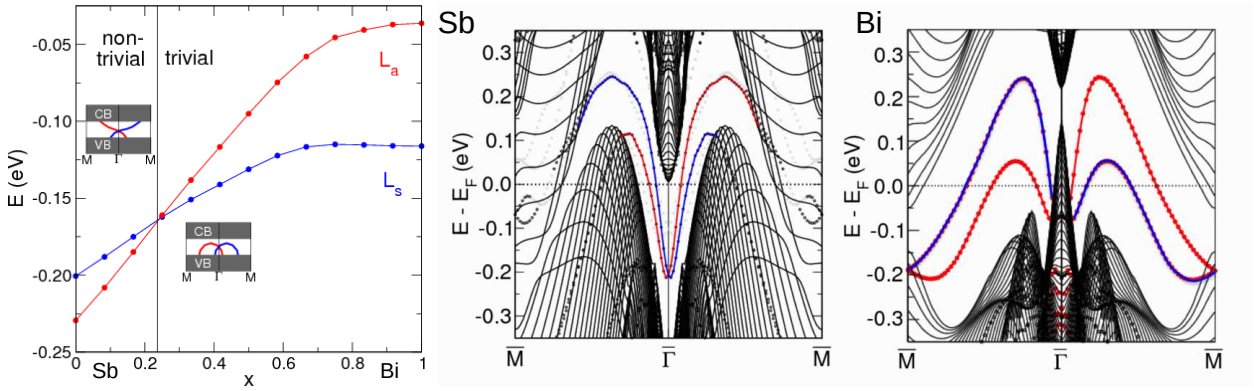


Fig. 11: *Left: Energetic position of the symmetric* (L_s) *and antisymmetric* (L_a) *eigenfunction at the* L *point in Bi as a function of spin-orbit coupling strength. To simulate a* Bi_xSb_{1-x} *alloy, the SOC strength is continuously varied between the values for Sb* ($x = 0$) *and Bi* ($x = 1$)*. Insets show schematically how the valence band (VB) and the conduction band (CB) are connected by a surface or edge state in the topologically non-trivial (left) and trivial (right) case. The calculated surface band structure of the Sb(111) and Bi(111) surface is shown in the middle and right panel, respectively. The surface states are shown in red and blue, indicating the spin-splitting due to the Rashba-Bychkov effect. Projected bulk-bandstructure is indicated by black lines outlining the projected bandgap where the surface states are observed.*

surface, spin-split surface states connect valence and conduction band and cross at the $\overline{\Gamma}$-point. On the Bi(111) surface, however, both branches of the surface state unite again in the valence band at the $\overline{\mathrm{M}}$ point, indicating that this surface state is not protected and can be removed e.g. by impurities at the surface. Note, that the $\overline{\Gamma}$ and the $\overline{\mathrm{M}}$ point are special electron momenta (so called time-reversal invariant momenta, TRIMs) where the electron is not moving and the splitting goes to zero. The band inversion at the (bulk) L point is also visible in the surface band structure at the $\overline{\mathrm{M}}$ point, which is the projection of the L point on the (111) plane.

It might be surprising that Sb, the lighter element, is a TI while the heavier Bi is topologically trivial as a bulk material – in a previous subsection we mentioned already that for (111) bilayers of these elements the situation is exactly opposite: the Bi (111) bilayer is a 2D-TI, the Sb bilayer is trivial. In both cases SOC drives the band-inversion, but the special band-topology of the 3D materials without SOC leads to different results.

Although the bandgap in the Bi_xSb_{1-x} alloy ($0.07 < x < 0.15$) is similar to that of HgTe/CdTe quantum wells, experimentally it is difficult to verify quantized conductance in these alloys. While HgTe/CdTe layers can be grown with high perfection, a similar control of the alloys seems difficult. Therefore, other TIs with larger bandgaps are more popular today, in particular A_2B_3 compounds with A = Bi, Sb and B = Te, Se. All combinations except for Sb_2Se_3 are TIs [47] with a single, Dirac-cone like edge state on the closed packed surfaces. We illustrate the situation for the Sb_2Te_3 (111) surface in fig. 12: Without SOC the compound has a 0.2 eV bandgap without surface states at the Fermi level, however at around -1.0 eV we see a parabolic surface state in the projected bulk bandgap. SOC leads to a band-inversion a the Γ point and a Dirac-cone shaped topologically protected surface state crosses the bandgap. The "trivial" surface state at -0.8 eV shows conventional Rashba-type spin-orbit splitting around Γ, but at larger k-vectors connects to different projected bulk bands [48]. This is a consequence of the opening of a spin-orbit gap in the bulk bandstructure around $k = 0.2$ Å^{-1} where, according to a general argument of Pendry and Gurman surface states have to be formed [49]. In a way this

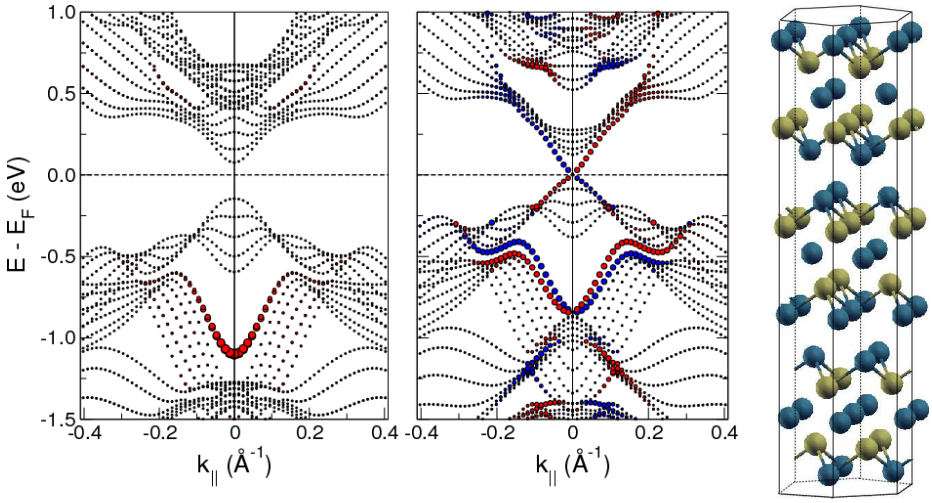


Fig. 12: *Bandstructure of a six quintuple-layer Sb_2Te_3 (111) film without (left) and with SOC (middle). The bulk structure with three QLs is shown on the right: blue spheres indicate Te atoms, yellow ones the Sb. The surface-localization of the states in the bandstructure without SOC is indicated by the size of the red circles, in case of the bandstructure with SOC, red/blue indicates also the spin-direction of the states.*

resembles the situation in graphene, where the edge-state connects two different band-edges at k and $-k$ in a spin-orbit gap.

Finally, we would like to mention that also HgTe, that was discussed previously as a 2D-TI, can be transformed into a 3D topological insulator: epitaxial strain allows to open a small bandgap in the otherwise semimetallic compound [50]. Very similar is the situation in α-Sn, which is naturally a semimetall with a bandstructure very similar to HgTe and also can be used as 3D-TI [51]. Some differences in the SOC effects and opening of a bandgap by strain come, however, from the fact that α-Sn crystallizes in a diamond structure with inversion symmetry, while this symmetry is missing in HgTe where e.g. the Dresselhaus effect can be seen.

5 Spin-orbit effects in magnetic systems

In all discussions above, spin-orbit coupling was considered as the dominant effect acting on the electrons spin. In absence of any other magnetic field, this is certainly justified and in heavy elements usually also external magnetic fields are too small to change the situation significantly. But if we consider the case of the bulk or a surface of a magnetic metal, like Gd(0001), certainly the exchange interaction will dominate and the effects change. This will be analyzed in the next subsection.

Furthermore, if SOC effects are of the order of the exchange interaction, they can give rise to new, anisotropic magnetic interactions. We will consider the case of the Dzyaloshinskii-Moriya interaction at surfaces as an example. Finally, we will consider the role of relativistic effects in the stabilization of the magnetic direction, i.e. the magnetic anisotropy. Also here, SOC plays an important (although not the only) role, especially in nanoscopic dimensions.

5.1 The Rashba-effect in an exchange field

We will discuss the effect of Rashba-type SOC and exchange interaction on a simple example, the surface state on the closed packed surface of Gd(0001). Exchange interaction splits this surface state into an occupied majority spin state and an unoccupied minority state. This splitting is mainly controlled by the $4f$ electrons of Gd and amounts to about 0.8 eV, which is large as compared to spin-orbit effects in this system. No matter how SOC affects the electrons of the surface state, their spin will remain more or less parallel to the exchange field, which is oriented in plane in the directions of nearest neighbor atoms by the magnetic anisotropy.
An electron traveling on the surface in a direction *perpendicular* to its spin quantization axis, will experience the potential gradient at the surface as a magnetic field parallel to its spin. Therefore, a magnetic coupling can arise and, neglecting the exchange splitting for the moment, the dispersion curves will split more or less similar to what is observed on a nonmagnetic surface. If, on the other hand, the propagation direction of the electron is parallel to its spin quantization axis, the field arising from SOC cannot couple to the electron's spin and no Rashba-like splitting can be observed.

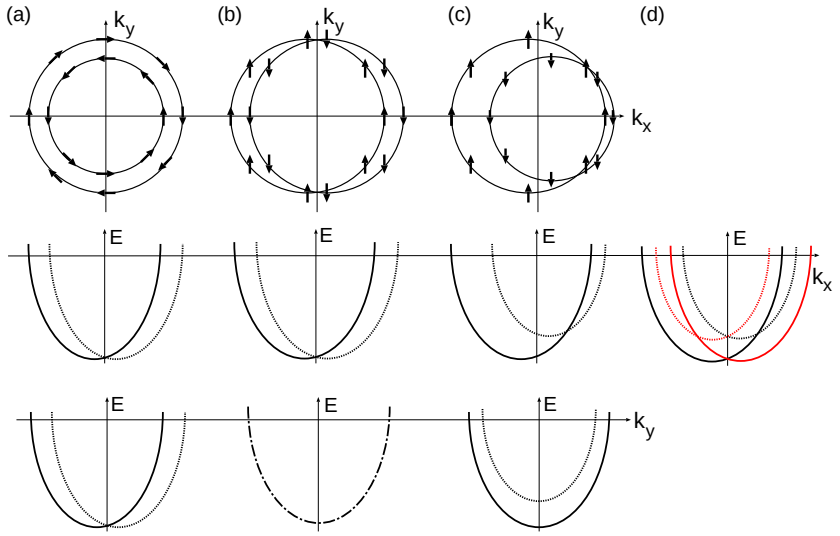


Fig. 13: *(a): Rashba splitting on a non-magnetic surface: the top panel shows the Fermi surface and the spin-polarization of the states at the Fermi level. In the middle and lower panel the bandstructures along two orthogonal directions in reciprocal space are shown. (b): The same relations as in (a) are shown, but now for the case where the spin-quantization axis of the electrons has been aligned in a particular direction. For electrons propagating in this direction, the Rashba splitting vanishes. (c): Effect of an additional exchange splitting on the situation as described in (b). (d): Combining the latter dispersion curves with the dispersion obtained from a sample with 180° rotated magnetization (red), the splitting of the states can be seen more clearly.*

Schematically, this situation is shown in figure 13. In contrast to the surface state on the non-magnetic surface, where the spin of the electron is always oriented perpendicular to the propagation direction and the surface normal, e_z, (with some deviation, depending on the shape of the potential [52]), on the spin-polarized surface, the spins are more or less collinear. This changes the shape of the Fermi surface significantly, especially if exchange splitting is considered (figure 13 (c)). If the exchange splitting is large, this leads to a Fermi surface consisting of a single

circle shifted away from the zone center. The consequences for the bandstructure are simple: along a certain direction in reciprocal space SOC will have no particular effect. In a direction orthogonal to this one, the dispersion curves for majority and minority spin will be shifted in opposite directions. For the eigenvalues this results in an expression

$$\varepsilon_{\downarrow(\uparrow)}(\mathbf{k}) = \varepsilon(\mathbf{k}) \pm IM \pm \alpha_{\mathrm{R}}(\mathbf{k} \times \mathbf{e}_z) \cdot \hat{\mathbf{M}} \tag{32}$$

where $\mathbf{M} = M\hat{\mathbf{M}}$ is the magnetization and IM represents the exchange splitting of the bands. Experimentally, for a single crystal surface, it is possible to measure with ARPES two spectra of the same surface, but rotated by 180 degrees. In the case of an in-plane anisotropy, as for Gd(0001), this rotation reverses the spin and leads, therefore, also to a picture as figure 13 (d). A comparison of these two spectra allowed to determine the Rashba splitting in Gd(0001), even though its magnitude is rather small [53].

5.2 Anisotropic exchange of adatoms on surfaces

In the last subsection, we assumed that the magnetic order at the surface is not influenced by spin-orbit coupling effects. If the exchange field is strong, all spins will align accordingly. On the other hand, if the exchange coupling is weak, spin-orbit coupling effects can substantially influence the magnetic interaction. The particular case of two distant impurities, which interact in a RKKY-type manner via a non-magnetic host which shows strong spin-orbit effects has been discussed by Smith [54]. He showed that the interaction between two magnetic atoms A and B (spins $\mathbf{S}_A$ and $\mathbf{S}_B$) via a non-magnetic third atom that acts on a conduction electron $\boldsymbol{\sigma}$ with a SOC term $\mathbf{l} \cdot \boldsymbol{\sigma}$ gives rise to an interaction $(\mathbf{S}_A \cdot \boldsymbol{\sigma})(\mathbf{l} \cdot \boldsymbol{\sigma})(\mathbf{S}_B \cdot \boldsymbol{\sigma})$. Taking the trace over the spin variable $\boldsymbol{\sigma}$ this term can be written as $(-i/4)\mathbf{l} \cdot (\mathbf{S}_A \times \mathbf{S}_B)$ and thus shows the form of the Dzyaloshinskii-Moriya (DM) interaction $\mathbf{D} \cdot (\mathbf{S}_A \times \mathbf{S}_B)$ that was derived from symmetry considerations by Dzyaloshinskii [55] and given a microscopical interpretation by Moriya [56] in the context of magnetic insulators showing weak ferromagnetism. Fert and Levy [57] derived an expression for this anisotropic exchange interaction of two magnetic atoms in spin-glasses doped with heavy impurity atoms which is of the form

$$\hat{H}_{\mathrm{DM}} = -V(\xi)\frac{\sin\left[k_{\mathrm{F}}(R_A + R_B + R_{AB}) + \eta\right]\hat{\mathbf{R}}_A \cdot \hat{\mathbf{R}}_B}{R_A R_B R_{AB}}(\hat{\mathbf{R}}_A \times \hat{\mathbf{R}}_B)(\mathbf{S}_A \times \mathbf{S}_B) \tag{33}$$

where $\mathbf{R}_A = R_A\hat{\mathbf{R}}_A$ and $\mathbf{R}_B = R_B\hat{\mathbf{R}}_B$ are the positions of the magnetic atoms measured from the nonmagnetic impurity and R_{AB} is the distance between the atoms A and B. $V(\xi)$ is a term that depends of the spin-orbit coupling constant of the nonmagnetic atom, ξ, k_{F} is the Fermi vector and η the phase shift induced by the impurity. The sinus term reflects the RKKY-type character of the interaction, while the two cross products determine the symmetry of the interaction.

This model can be translated to the case of two magnetic atoms on a surface, where the magnetic interaction is mediated by surface states which show strong SOC effects. Such a situation might be imagined, if e.g. two Mn atoms are placed on a Bi surface (figure 14). If the easy magnetic axis is out-of-plane, a slight tilting of the magnetic moments results in a finite value for $\mathbf{S}_A \times \mathbf{S}_B$ which is then parallel to $\hat{\mathbf{R}}_A \times \hat{\mathbf{R}}_B$ and leads to a non-vanishing contribution of H_{DM} (equation (33)). If the easy magnetic axis is in-plane (right of figure 14) and the surface normal is the hard axis, a small tilting of the magnetic moments results in a vector $\mathbf{S}_A \times \mathbf{S}_B$ that is normal to $\hat{\mathbf{R}}_A \times \hat{\mathbf{R}}_B$ and equation (33) will give no contribution to the total energy. Of course,

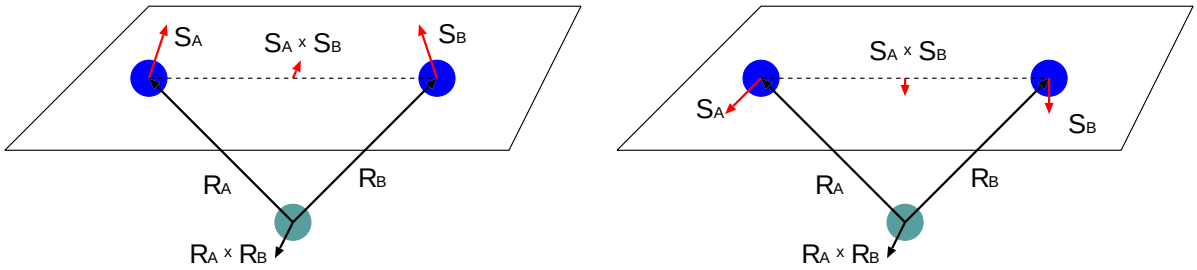


Fig. 14: *Two magnetic adatoms (A, B) on a surface interacting with a surface atom at the center. The distance between the surface atoms and the adatoms is $\mathbf{R}_A$ and $\mathbf{R}_B$. The spins of the adatoms are almost perpendicular to the surface (left) or in the surface plane (right), but sightly canted to give a finite value for $\mathbf{S}_A \times \mathbf{S}_B$*

on a surface the scattering will involve all surface atoms and in general it will depend not only on the direction of the spins of the adatoms, but also on the symmetry of the surface whether a DM interaction will occur for a specific arrangement of the spins. This will be discussed later in more detail.
If we extend the two impurities in figure 14 to a chain of magnetic atoms, where the spins of two neighboring atoms, i and j, are canted slightly, an interaction of this kind

$$\hat{H}_{\mathrm{DM}} = \mathbf{D}_{ij} \cdot (\mathbf{S}_i \times \mathbf{S}_j) \tag{34}$$

will favor spin-spiral structure. Since the DM interaction has to compete with the Heisenberg-type (symmetric) exchange interaction, these structures will probably be of long wavelength. Such long-ranged magnetic structures can be found on surfaces [58] or in domain walls [59] of thin magnetic films. It is here of particular importance, that the DM interaction gives these spiraling magnetic structures a unique sense of rotation. If e.g. a magnetic domain in a thin film is bounded by domain walls that rotate in the same direction, it is harder to expel the domain with a magnetic field [60].
In two-dimensionally modulated structures, e.g. superpositions of spin-spirals (multi $\mathbf{q}$ states) the DM interaction can also give rise to interesting effects: as it selects a specific winding sense of the magnetic structure, spin-textures can appear that are of the same form as the skyrmions and merons we encountered already in figure 10. Also here topology has a stabilizing effect, as found e.g. for an Fe monolayer on Ir(111) [61]. A more detailed discussion can be found in the contribution on complex magnetism (C4).

5.3 The magnetocrystalline anisotropy

As we discussed in the introduction, the control over the spatial orientation of spins or the magnetization in a crystal is a delicate, but important subject. Generally, a dependence of the total energy of the magnetization-direction (with respect to the crystal axes) is termed a magnetic anisotropy. This anisotropy fixes the magnetization direction w.r.t. the lattice and allows for a stable magnetization direction in a material, which is the basis of almost all magnetic applications.
The free energy F of a system with a magnetization $\mathbf{M}$ can be expanded in terms of polar and azimuthal angles, θ and ϕ, of the magnetization direction. This can be done either in an expansion in spherical harmonics, or, usually, in terms of the directional cosines $(\alpha_1, \alpha_2, \alpha_3) =$

$(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ of $\mathbf{M}$ with respect to the crystal axes. For example in a cubic lattice, the first terms of this expansion read:

$$\begin{aligned} F(\mathbf{M}) &= K_0 + K_1(\alpha_1^2\alpha_2^2 + \alpha_1^2\alpha_3^2 + \alpha_2^2\alpha_3^2) + K_2\alpha_1^2\alpha_2^2\alpha_3^2 + \ldots \\ &= K_0 + \frac{K_1}{64}\{(3 - 4\cos 2\theta + \cos 4\theta)(1 - \cos 4\phi) + 8(1 - \cos 4\theta)\} + \\ &\quad + \frac{K_2}{256}(1 - \cos 2\theta - 2\cos 4\theta + \cos 6\theta)(1 - \cos 4\phi) + \ldots \end{aligned} \tag{35}$$

In contrast to cubic systems, that show no single crystal axis that is non-equivalent to other axes, tetragonal or hexagonal lattices have a unique high symmetry axis (in these cases the c-direction). In these uniaxial systems the expansion of F is given by

$$F(\mathbf{M}) = K_0 + K_1\sin^2\theta + K_2\sin^4\theta + \ldots \tag{36}$$

Anisotropy constants for several materials are collected in table 1. We see, that cubic materials generally have smaller K's than uniaxial systems. This is understandable, since for cubic systems the anisotropy constants refer to higher order expansions in the α's than for uniaxial systems.

Table 1: *Magnetic anisotropy constants K_1 and K_2 for some cubic and uniaxial magnetic materials in J/m^3 at $T = 4.2$K (Ref. [62])*

Class	Material	K_1	K_2	Symmetry
$3d$ Metals	Fe	$5.2 \cdot 10^4$	$-1.8 \cdot 10^4$	cub.
	Co	$7.0 \cdot 10^5$	$1.8 \cdot 10^5$	uni.
	Ni	$-1.2 \cdot 10^5$	$3.0 \cdot 10^4$	cub.
$4f$ Metals	Gd	$-1.2 \cdot 10^5$	$8.0 \cdot 10^4$	uni.
	Tb	$-5.7 \cdot 10^7$	$-4.6 \cdot 10^6$	uni.
	Er	$1.2 \cdot 10^7$	$-3.9 \cdot 10^6$	uni.
Hard Magnets	Sm Co_5	$7 \cdot 10^6$	—	uni.
	Nd Co_5	$-4 \cdot 10^7$	—	uni.
Spinel Ferrites	Fe_3 O_4	$-2 \cdot 10^4$	$-9 \cdot 10^3$	cub.
	Co Fe_2 O_4	$\approx 10^6$	—	cub.

There are several interactions that can lead to a magnetic anisotropy, e.g. the dipolar interaction [63] which is of substantial importance in bulk materials. In low-dimensional magnets also other interactions that lead to magnetically anisotropic behavior can become dominant, most importantly the spin-orbit coupling. We will call the contribution that results from the anisotropy of the spin-orbit interaction here the magnetocrystalline anisotropy (MCA). It should be noticed, however, that this term is also sometimes used in context of the contributions from the dipole-dipole interaction to discriminate "continuum terms" from terms that arise due to the discrete nature of the crystal lattice.

As mentioned shortly at the end of subsection 5.1, in a magnetic system the spin-orbit induced splittings in a bandstructure will be influenced by the direction of the spin-quantization axis (SQA). For example, if the easy axis is perpendicular to the surface plane, the effect of the

Rashba-field for all surface state electrons is perpendicular to the spin direction and no additional effect is expected. This dependence on the SQA gives a small, but important contribution to the total energy of a magnetic system since in a non-relativistic Hamiltonian there is no term which could give a dependence of the total energy on the direction of the SQA.
If we consider a band-crossing near the Fermi level that occurs without SOC taken into account, depending on the magnetization direction SOC might open up a gap at this crossing or not. In the former case, it is obviously possible to gain energy that adds to the MAE. Whether a gap opens or not depends now on the symmetry of the states that cross. Let us analyze the situation for d-orbitals: Evidently, out of certain d-levels only orbital moments pointing in a certain direction can be formed. E.g. a d_{xy} and a $d_{x^2-y^2}$ orbital can only be combined to form an orbital moment in z direction. An orbital moment pointing in x-direction has to be formed from electrons that can move in the (y,z)-plane, and this is impossible within only the d_{xy} and $d_{x^2-y^2}$ orbitals. If now two appropriate orbitals are degenerate and occupied by a single electron (and thus forming the Fermi level), it is rather straightforward to identify the resulting direction of the orbital moment using group theory [64]. In a metal, where several bands are crossing the Fermi level, ϵ_{F}, it is basically the sum of all contributions from bands near ϵ_{F} that determine the orbital moment. In second-order perturbation theory the expectation value of the orbital moment operator $\mathbf{L}$ can be written as:

$$\langle \mathbf{L} \rangle = \sum_{i,j} \frac{\langle \psi_i | \mathbf{L} | \psi_j \rangle \left\langle \psi_j | \hat{H}_{\mathrm{SOC}} | \psi_i \right\rangle}{\epsilon_i - \epsilon_j} f(\epsilon_i) \left[1 - f(\epsilon_j) \right], \tag{37}$$

where H_{so} is the spin-orbit coupling Hamiltonian and f is the Fermi function ensuring that the wavefunction ψ_i is occupied and ψ_j is unoccupied. Van der Laan [64] has shown, that in the absence of spin-flip terms (i.e. when the majority and minority band are well separated by the exchange interaction), the spin-orbit coupling changes the total energy of a system in second-order perturbation theory as:

$$\delta E = \sum_{i,j} \frac{\left\langle \psi_i | \hat{H}_{\mathrm{SOC}} | \psi_j \right\rangle \left\langle \psi_j | \hat{H}_{\mathrm{SOC}} | \psi_i \right\rangle}{\epsilon_i - \epsilon_j} f(\epsilon_i) \left[1 - f(\epsilon_j) \right] \approx -\frac{1}{4} \xi \hat{\mathbf{S}} \cdot \left[\left\langle \mathbf{L}^{\downarrow} \right\rangle - \left\langle \mathbf{L}^{\uparrow} \right\rangle \right] \tag{38}$$

where ξ is the radial part of the spin-orbit Hamiltonian (eq. 8), $\hat{\mathbf{S}}$ is the direction of the spin moment, and $\mathbf{L}^{\downarrow}$ and $\mathbf{L}^{\uparrow}$ are the orbital moment vectors of the spin-down and spin-up bands, respectively. If the spin-up band is completely filled, we see that energy change, δE, is proportional to the size of the orbital moment and the magnetocrystalline anisotropy energy (MAE), i.e. the difference of δE for two different magnetization directions, will be proportional to the difference in the orbital moments. This relation between orbital moment anisotropy and MAE was first derived by Bruno [65].
A drastic example is provided by a organometallic molecule, the $\mathrm{Eu_2(C_8H_8)_3^{-1}}$) anion (fig. 15). If the magnetization is along the axis of the molecule, the highest occupied and lowest unoccupied level can couple to form a large orbital moment ($0.19\mu_{\mathrm{B}}$ in the DFT calculation) and the splitting is large. Since there is a single electron in the highest occupied level, the energy gain is large. On the other hand, if the magnetization is along the radial direction the states can not couple and the splitting is small. The energy difference for the two spin-directions is 13.7 eV in favor of the radial direction. In contrast, in the neutral molecule this level is unoccupied and the easy axis is in radial direction [66]

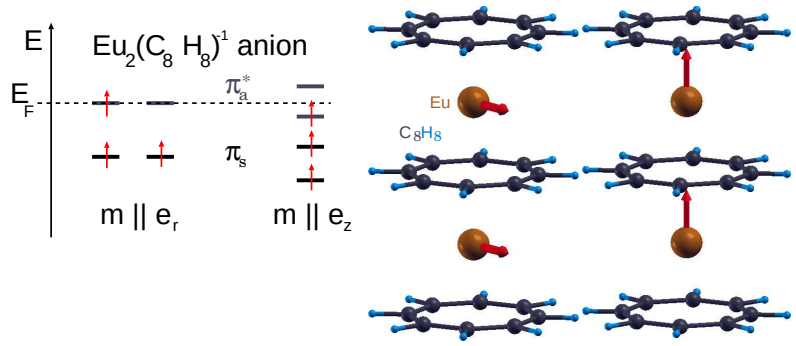


Fig. 15: *Left: Highest occupied and lowest unoccupied levels in an* $\mathrm{Eu_2(C_8H_8)_3^{-1}}$ *anion for magnetization is in radial direction* $(m\|e_r)$ *and along the molecule axis* $(m\|e_z)$ *as schematically shown in the middle and right panels, respectively.*

Practically, in a DFT calculation the magnetocrystalline anisotropy energy (MAE) can be obtained from the difference of total energies from Hamiltonians including the spin-orbit coupling term with the magnetization pointing in two different directions. Practically, one starts from a solution Ψ_0 of the Schrödinger equations (possibly including scalar-relativistic corrections, cf. section 2.1), and then solves the Hamiltonian including the spin-orbit coupling term with the spin-quantization axis turned into the required direction by means of a spin-rotation matrix U:

$$\begin{aligned}\left\langle U\Psi_0|\hat{H}_\mathrm{S}+\xi\boldsymbol{\sigma}\cdot\mathbf{L}|U\Psi_0\right\rangle &= \left\langle\Psi_0|\hat{H}_\mathrm{S}|\Psi_0\right\rangle+\xi\left\langle U\Psi_0|\boldsymbol{\sigma}\cdot\mathbf{L}|U\Psi_0\right\rangle = \\ &= \varepsilon_0+\xi\left\langle\begin{pmatrix}\Psi_0^\uparrow\\ \Psi_0^\downarrow\end{pmatrix}\middle|U^\dagger\begin{pmatrix}L_z & L_x-iL_y\\ L_x+iL_y & -L_z\end{pmatrix}U\middle|\begin{pmatrix}\Psi_0^\uparrow\\ \Psi_0^\downarrow\end{pmatrix}\right\rangle \quad (39)\end{aligned}$$

If ξ or the orbital moment is small, the last part of eq. 39 is only a small correction to the energy ε_0 obtained from the Schrödinger equation and the magnetization direction of the solution will point into the direction of the spin-quantization axis described by U.

From the above equations it is clear that both, strong spin-orbit coupling and a sizable orbital moment, $\mathbf{L}$, are necessary for a large contribution to the magnetic anisotropy. But it is also necessary that the spin-orbit interaction gives different energy contributions for different magnetizations of the sample. In principle there are two possibilities to imagine how this could happen: (i) the orbital moment is fixed to the lattice and its projection on the axis of the spin moment varies with the magnetization direction or (ii) the spin and orbital moments are collinear and depending on the magnetization direction the size of the orbital moment varies. Normally, we observe collinear spin- and orbital moments. The rotation of the orbital moment by an external magnetic field can then lead to structural changes of the crystal. This phenomenon is called magnetostriction and is discussed e.g. in ref. [63].

The MAE is a typically a small energy, for elemental bulk magnets it is in the order of microelectronvolts (μeV). This is mainly a consequence of the high symmetry in these bulk systems. Low-dimensional systems (thin films, chains and wires) can show much higher MAE's, up to a few milli-electronvolts. Since other sources of magnetic anisotropy can be even smaller in these systems, spin-orbit coupling can get very important in magnetic nanostructures.

Sometimes, in analogy to the Heisenberg Hamiltonian describing the exchange interaction in a crystal, the spin-orbit coupling is cast into a form $\sum_i \xi_i L_i S_i$ where i is a particular atomic

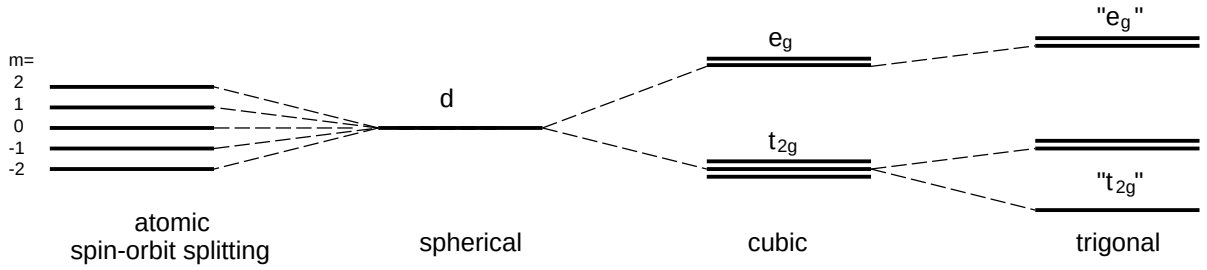


Fig. 16: *Splitting of the d levels in an atom (left) and a solid under the influence of a cubic (octahedral) crystal field (middle). A further small deformation of the lattice (e.g. as a result of bonding in the spinels) leads to an additional splitting (right). Single occupancy of the remaining doubly degenerate levels can lead to the formation of an orbital moment.*

site. Then, evidently, another term coupling the spin of a site i to the orbital motion at site j is conceivable: $C_{ij}L_jS_i$. This spin-other orbit interaction is, like the dipole-dipole interaction, derived from the Breit equation. In the Hartree approximation it was included in *ab-initio* calculations but was found to be much weaker than the formerly described spin- (same) orbit interaction [67].
Other sources of magnetic anisotropy that come from the interaction on two sites can be derived from SOC in a way similar to our introduction of the Dzyaloshinskii-Moriya interaction: Higher order processes of the spin-orbit scattering of conduction electrons lead to (pseudo) dipolar interactions that introduce an anisotropy in the system [54]. Generally, these higher order terms are very short ranged and mainly discussed in magnetic insulators where, due to the lack of a Fermi surface, the mechanism discussed above are not effective.

5.4 Orbital moments and single-ion anisotropy

In a bulk crystal of high symmetry (e.g. fcc or bcc), most orbitals of the valence electrons are directed in the nearest or next-nearest neighbor directions. The crystal environment of the atom confines the electrons then to directional bonds and effectively suppresses thereby orbital motion. Spin-orbit coupling tries to counteract this "quenching" of the orbital moments and – in some cases – a sizable orbital moment can occur even in a bulk crystal.
As a prototypical example let us consider magnetite, Fe_3O_4, a classical magnetic material crystallizing in the inverse spinel structure. The O^{2-} ions form a close packed cubic lattice and $1/8$ of the tetrahedral sites are occupied with Fe^{3+}, while the remaining Fe^{2+} and Fe^{3+} ions each occupy $1/4$ of the octahedral sites. In these crystal fields, the Fe d-levels will split up in t_{2g} and e_g levels. In the tetrahedral crystal field, the e_g levels will have lower energy, while on the octahedral sites the t_{2g} levels will be lower in energy. While in the free atom the orbitals could be classified according to their magnetic quantum number $m = -2, -1, 0, 1, 2$, the crystal field now forces the formation of new linear combinations to form orbitals with vanishing orbital moment (e.g. the t_{2g} states d_{yz} and d_{zx} are linear combinations of the atomic $m = -1$ and $m = +1$ orbitals). Due to a small trigonal deformation, the t_{2g} levels show an additional splitting in the spinel structure (cf. figure 16). Assuming that the Fe^{3+} ions have a d^5 configuration and this half filled d shell corresponds to a zero orbital moment, these atoms will not contribute much to the magnetocrystalline anisotropy. In Fe^{2+}, the additional electron will occupy the lowest, split-of level of the t_{2g} states. When we substitute Fe by Co, we have another d-electron

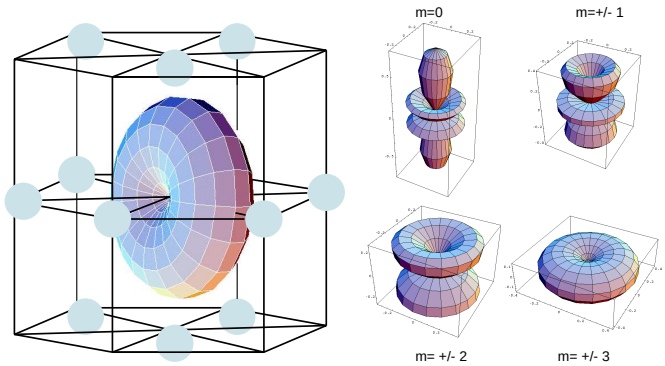


Fig. 17: *Left: hcp crystal with a $4f$ wavefunction with $|m| = 3$. As in bulk Tb, the easy axis is in the close-packed planes. The absolute values of some other f wavefunctions are shown on the right.*

that can now occupy the remaining, doubly degenerate t_{2g}-like states. Since this electron is not involved in bonding, these states can now again form linear combinations with considerable orbital moments in a specific direction. As has been shown by Slonczewski [68], it is this orbital moment that leads to the pronounced difference in the magnetic anisotropy between magnetite and $CoFe_2O_4$ (cf. table 1).

For a given system it is often hard to tell which kind of effect, the dipole - dipole or the spin-orbit interaction, will give the major contribution to the anisotropy energy. For systems with very small orbital moment, like bulk Eu that has a half-filled $4f$ shell ($S = 7/2, L = 0$) and only s-electrons in the conduction band, the spin-orbit interaction is weak (as can be inferred already from the fact that bcc Eu has a spin-spiral magnetic ground state). Therefore, the shape anisotropy is likely to dominate except for spherical or elliptical samples that can be easily magnetized in all different directions. Gd, too, has a half-filled $4f$ shell, but one d-electron more than Eu. Its magnetic anisotropy is also small, but due to the additional d-electron the ordering temperature is much higher than in Eu. With again one more electron, hcp Tb has an orbital moment of $L = 3$ and there is no easy way to change the magnetization direction of a Tb crystal except for heating it up beyond the Curie-temperature and cooling under an external field in a symmetry-equivalent direction. In the compressed hcp lattice of Tb ($c/a = 1.59$) the flat, pancake-like minority $4f$ electron is locked in the crystal lattice [69] and the easy axis is parallel to the b-axis (fig. 17. This strong anisotropy, that can be regarded as coming from a single atom is often referred to as single-ion anisotropy.

From the above discussion it should be clear, that orbital effects are most efficiently quenched in a bulk-like environment. There, the coordination is highest and high symmetry can additionally reduce the magnetic anisotropy. But in (quasi) low-dimensional systems like thin films, chains and wires, or adatoms on a substrate spin-orbit coupling will be of much higher importance and large anisotropies can be expected.

Acknowledgment: I gratefully thank Ph. Hofmann, E. V. Chulkov, the late K. Starke and all their co-workers who ignited my interest in spin-orbit phenomena and Rashba-Bychkov type effects at surfaces, S. Murakami who introduced me to the topological insulators, and S. Heinze with whom we ventured in the field of magnetic skyrmions. Finally my thanks go to our local people like Y. Mokrousov, D. Wortmann, Ph. Mavropoulos and S. Blügel and all who share my excitement for relativistic effects in solids.

References

[1] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems*, Plenum, New York, 1977.

[2] L. H. Thomas, Motion of the spinning electron, Nature **117**, 514 (1926).

[3] M. Chrysos, The non-intuitive $\frac{1}{2}$ Thomas factor: a heuristic argument with classical electromagnetism, Eur. J. Phys **27**, 1 (2006).

[4] F. Hund, Zur Deutung verwickelter Spektren, insbesondere der Elemente Scandium bis Nickel, Z. Physik **33**, 345 (1929).

[5] A. R. Mackintosh and O. K. Andersen, The electronic structure of transition metals, in *Electrons at the Fermi Surface*, edited by M. Springford, page 149, London, 1980, Cambridge University Press.

[6] S. Krupička, *Physik der Ferrite und der verwandten magnetischen Oxide*, Vieweg, Braunschweig, 1973.

[7] G. Dresselhaus, Spin-orbit coupling effects in zinc blende structures, Phys. Rev. **100**, 580 (1955).

[8] R. Winkler, SpinOrbit Coupling Effects in Two-Dimensional Electron and Hole Systems, in *Springer Tracts on Modern Physics*, volume 191, Berlin Heidelberg, 2003, Springer.

[9] E. I. Rashba, Symmetry of bands in wurzite-type crystals. 1. Symmetry of bands disregarding spin-orbit interaction, Sov. Phys. Solid State **1**, 368 (1959).

[10] E. I. Rashba and V. I. Sheka, Simmetriya energeticheskikh zon v kristallakh tipa vyurtsita. II. Simmetriya zon s uchyotom spinovykh vzaimodeistvii, Fizika tverd. tela **1**, (2), 162 (1959).

[11] K. Ishizaka, M. S. Bahramy, H. Murakawa, M. Sakano, T. Shimojima, T. Sonobe, K. Koizumi, S. Shin, H. Miyahara, A. Kimura, K. Miyamoto, T. Okuda, H. Namatame, M. Taniguchi, R. Arita, N. Nagaosa, K. Kobayashi, Y. Murakami, R. Kumai, Y. Kaneko, Y. Onose, and Y. Tokura, Giant Rashba-type spin splitting in bulk BiTeI, Nature Mater. **10**, 521 (2011).

[12] Y. A. Bychkov and E. I. Rashba, Oscillatory effects and the magnetic-susceptibility of carriers in inversion-layers, J. Phys. C: Solid State Phys. **17**, 6039 (1984).

[13] Y. A. Bychkov and E. I. Rashba, Properties of a 2D eelctron-gas with lifted spectral degeneracy, Sov. Phys. JETP Lett **39**, 78 (1984).

[14] T. Schäpers, J. Knobbe, and V. A. Guzenko, Effect of Rashba spin-orbit coupling on magnetotransport in InGaAs/InP quantum wire structures, Phys. Rev. B **69**, 235323 (2004).

[15] S. LaShell, B. A. McDougall, and E. Jensen, Spin Splitting of an Au(111) Surface State Band Observed with Angle Resolved Photoelectron Spectroscopy, Phys. Rev. Lett. **77**, 3419 (1996).

[16] G. Bihlmayer, Y. M. Koroteev, P. M. Echenique, E. V. Chulkov, and S. Blügel, The Rashba-effect at metallic surfaces, Surf. Sci. **600**, 3888 (2006).

[17] S. Datta and B. Das, Electronic analog of the electro-optic modulator, Appl. Phys. Lett. **56**, 665 (1990).

[18] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, Quantized Hall Conductance in a Two-Dimensional Periodic Potential, Phys. Rev. Lett. **49**, 405 (1982).

[19] L. Fu and C. L. Kane, Topological insulators with inversion symmetry, Phys. Rev. B **76**, 045302 (2007).

[20] J. Li, I. Martin, M. Büttiker, and A. F. Morpurgo, Marginal topological properties of graphene: a comparison with topological insulators, Phys. Scr. **T146**, 014021 (2012).

[21] D. P. DiVincenzo and E. J. Mele, Self-consistent effective-mass theory for intralayer screening in graphite intercalation compounds, Phys. Rev. B **29**, 1685 (1984).

[22] C. L. Kane and E. J. Mele, Quantum Spin Hall Effect in Graphene, Phys. Rev. Lett. **95**, 226801 (2005).

[23] M. Gmitra, S. Konschuh, C. Ertler, C. Ambrosch-Draxl, and J. Fabian, Band-structure topologies of graphene: Spin-orbit coupling effects from first principles, Phys. Rev. B **80**, 235431 (2009).

[24] S. Konschuh, M. Gmitra, and J. Fabian, Tight-binding theory of the spin-orbit coupling in graphene, Phys. Rev. B **82**, 245412 (2010).

[25] F. D. M. Haldane, Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the "Parity Anomaly", Phys. Rev. Lett. **61**, 2015 (1988).

[26] X.-L. Qi and S. S. Zhang, The quantum spin Hall effect and topological insulators, Physics Today **63**, 33 (2010).

[27] M. Fujita, K. Wakabayashi, K. Nakada, and K. Kusakabe, Peculiar Localized State at Zigzag Graphite Edge, J. Phys. Soc. Jpn. **65**, 1920 (1996).

[28] J. D. Sau, R. M. Lutchyn, S. Tewari, and S. D. Sarma, Generic New Platform for Topological Quantum Computation Using Semiconductor Heterostructures, Phys. Rev. Lett. **104**, 040502 (2010).

[29] A. Varykhalov, J. Sánchez-Barriga, A. M. Shikin, C. Biswas, E. Vescovo, A. Rybkin, D. Marchenko, and O. Rader, Electronic and magnetic properties of quasifreestanding graphene on Ni, Phys. Rev. Lett. **101**, 157601 (2008).

[30] D. Marchenko, A. Varykhalov, M. R. Scholz, G. Bihlmayer, E. I. Rashba, A. Rybkin, A. M. Shikin, and O. Rader, Giant Rashba splitting in graphene due to hybridization with gold, Nature Comm. **3**, 1232 (2012).

[31] D. Marchenko, A. Varykhalov, M. R. Scholz, J. Sánchez-Barriga, O. Rader, A. Rybkina, A. M. Shikin, T. Seyller, and G. Bihlmayer, Spin-resolved photoemission and ab initio theory of graphene/SiC, Phys. Rev. B **88**, 075422 (2013).

[32] C. Weeks, J. Hu, J. Alicea, M. Franz, and R. Wu, Engineering a robust quantum spin Hall state in graphene via adatom deposition, Phys. Rev. X **1**, 021001 (2011).

[33] S. Murakami, Quantum Spin Hall Effect and Enhanced Magnetic Response by Spin-Orbit Coupling, Phys. Rev. Lett. **97**, 236805 (2006).

[34] M. Wada, S. Murakami, F. Freimuth, and G. Bihlmayer, Localized edge states in two-dimensional topological insulators: Ultrathin Bi films, Phys. Rev. B **83**, 121310(R) (2011).

[35] T. Hirahara, G. Bihlmayer, Y. Sakamoto, M. Yamada, H. Miyazaki, S. i. Kimura, S. Blügel, and S. Hasegawa, Interfacing 2D and 3D topological insulators: Bi(111) bilayer on Bi_2Te_3, Phys. Rev. Lett. **107**, 166801 (2011).

[36] H. Zhang, F. Freimuth, G. Bihlmayer, M. Lezaic, S. Blügel, and Y. Mokrousov, Engineering quantum anomalous Hall phases with orbital and spin degrees of freedom, Phys. Rev. B **87**, 205132 (2013).

[37] B. A. Bernevig, T. L. Hughes, and S.-C. Zhang, Quantum Spin Hall Effect and Topological Phase Transition in HgTe Quantum Wells, Science **314**, 1757 (2006).

[38] M. König, S. Wiedmann, C. Brüne, A. Roth, H. Buhmann, L. W. Molenkamp, X.-L. Qi, and S.-C. Zhang, Quantum Spin Hall Insulator State in HgTe Quantum Wells, Science **318**, 766 (2007).

[39] E. O. Kane, Band Structure of Indium Antimonide, J. Phys. Chem. Solids **1**, 249 (1957).

[40] J. M. Luttinger and W. Kohn, Motion of Electrons and Holes in Perturbed Periodic Fields, Phys. Rev. **97**, 869 (1955).

[41] N. A. Cade and P. M. Lee, Self consistent energy band structures for HgTe and CdTe, Solid State Comm. **56**, 637 (1985).

[42] N. E. Christensen, I. Gorczyca, O. B. Christensen, U. Schmid, and M. Cardona, Band structure and heterojunctions of II-VI materials, J. Cryst. Growth **101**, 318 (1990).

[43] J. E. Avron, D. Osadchy, and R. Seiler, A Topological Look at the Quantum Hall Effect, Physics Today **August 2003**, 38 (2003).

[44] G. E. Volovik, The Universe in a Helium Droplet, in *The International Series of Monographs on Physics*, edited by J. Birman and S. F. Edwards, volume 117, Oxford, 2003, Clarendon Press.

[45] T. Hirahara, K. Miyamoto, A. Kimura, Y. Niinuma, G. Bihlmayer, E. V. Chulkov, T. Nagao, I. Matsuda, S. Qiao, K. Shimada, H. Namatame, M. Taniguchi, and S. Hasegawa, Origin of the surface-state band-splitting in ultrathin Bi films: From a Rashba effect to a parity effect, New J. Phys. **10**, 083038 (2008).

[46] D. Hsieh, Y. Xia, L. Wray, D. Qian, A. Pal, J. H. Dil, J. Osterwalder, F. Meier, G. Bihlmayer, C. L. Kane, Y. S. Hor, R. J. Cava, and M. Z. Hasan, Observation of unconventional quantum spin textures in topological insulators, Science **323**, 919 (2009).

[47] H. Zhang, C.-X. Liu, X.-L. Qi, X. Dai, Z. Fang, and S.-C. Zhang, Topological insulators in Bi_2Se_3, Bi_2Te_3 and Sb_2Te_3 with a single Dirac cone on the surface, Nature Phys. **5**, 438 (2009).

[48] C. Pauly, G. Bihlmayer, M. Liebmann, M. Grob, A. Georgi, D. Subramaniam, M. R. Scholz, J. Sànchez-Barriga, A. Varykhalov, S. Blügel, O. Rader, and M. Morgenstern, Probing two topological surface bands of Sb_2Te_3 by spin-polarized photoemission spectroscopy, Phys. Rev. B **86**, 235106 (2012).

[49] J. B. Pendry and S. J. Gurman, Theory of surface states: General criteria for their existence, Surf. Sci. **49**, 87 (1975).

[50] C. Brüne, C. X. Liu, E. G. Novik, E. M. Hankiewicz, H. Buhmann, Y. L. Chen, X. L. Qi, Z. X. Shen, S. C. Zhang, and L. W. Molenkamp, Quantum Hall Effect from the Topological Surface States of Strained Bulk HgTe, Phys. Rev. Lett. **106**, 126803 (2011).

[51] A. Barfuss, L. Dudy, M. R. Scholz, H. Roth, P. Höpfner, C. Blumenstein, G. Landolt, J. H. Dil, N. C. Plumb, M. Radovic, A. Bostwick, E. Rotenberg, A. Fleszar, G. Bihlmayer, D. Wortmann, W. Hanke, R. Claessen, and J.Schäfer, Elemental topological insulator with a tunable Fermi level: Strained α-Sn on InSb(001), Phys. Rev. Lett. **111**, 157205 (2013).

[52] J. Henk, A. Ernst, and P. Bruno, Spin polarization of the *L*-gap surface states on Au(111), Phys. Rev. B **68**, 165416 (2003).

[53] O. Krupin, G. Bihlmayer, K. Starke, S. Gorovikov, J. E. Prieto, K. Döbrich, S. Blügel, and G. Kaindl, Rashba effect at magnetic metal surfaces, Phys. Rev. B **71**, 201403 (R) (2005).

[54] D. A. Smith, New mechanisms for magnetic anisotropy in localised s-state moment materials, J. Magn. Magn. Mater. **1**, 214 (1976).

[55] I. Dzyaloshinsky, Thermodynamic Theory of "Weak" Ferromagnetism In Antiferromagnetic Substances, J. Phys. Chem. Solids **4**, 241 (1958).

[56] T. Moriya, Anisotropic Superexchange Interaction and Weak Ferromagnetism, Phys. Rev. **120**, 91–98 (1960).

[57] A. Fert and P. M. Levy, Role of Ansiotropic Exchange Interactions in Determining the Properties of Spin-Glasses, Phys. Rev. Lett. **44**, 1538 (1980).

[58] M. Bode, M. Heide, K. v. Bergmann, P. Ferriani, S. Heinze, G. Bihlmayer, A. Kubetzka, O. Pietzsch, S. Blügel, and R. Wiesendanger, Chiral magnetic order at surfaces driven by inversion asymmetry, Nature **447**, 190 (2007).

[59] M. Heide, G. Bihlmayer, and S. Blügel, Dzyaloshinskii-Moriya interaction accounting for the orientation of magnetic domains in ultrathin films: Fe/W(110), Phys. Rev. B **78**, 140403(R) (2008).

[60] S. Heinze, K. v. Bergmann, and G. Bihlmayer, Gitter aus magnetischen Wirbeln, Physik in unserer Zeit **43**, 6 (2012).

[61] S. Heinze, K. v. Bergmann, M. Menzel, J. Brede, A. Kubetzka, R. Wiesendanger, G. Bihlmayer, and S. Blügel, Spontaneous atomic-scale magnetic skyrmion lattice in two dimensions, Nature Phys. **7**, 713 (2011).

[62] R. O. Handley, *Modern magnetic materials*, Wiley, New York, 2000.

[63] S. Blügel, *Materie und Material (2)*, chapter Magnetische Anisotropie und Magnetostriktion, Schriften des FZ-Jülich, Jülich, 1999.

[64] G. van der Laan, Microscopic origin of magnetocrystalline anisotropy in transition metal thin films, J.Phys.: Condens. Matter **10**, 3239 (1998).

[65] P. Bruno, Tight-binding approach to the orbital magnetic moment and magnetocrystalline anisotropy of transition-metal monolayers, Phys. Rev. B **39**, 865 (1989).

[66] N. Atodiresei, P. H. Dederichs, Y. Mokrousov, L. Bergqvist, G. Bihlmayer, and S. Blügel, Controlling the magnetization direction in molecules via their oxidation state, Phys. Rev. Lett. **100**, 117207 (2008).

[67] M. D. Stiles, S. V. Halilov, R. A. Hyman, and A. Zangwill, Spin-other-orbit interaction and magnetocrystalline anisotropy, Phys. Rev. B **64**, 104430 (2001).

[68] J. C. Slonczewski, Magnetic Annealing, in *Magnetism*, edited by G. Rado and H. Suhl, volume 1, page 205, New York, 1963, Academic Press.

[69] S. Chikazumi, *Physics of Ferromagnetism*, Oxford University Press, 1964.

B 1 Model Hamiltonians and Second Quantization [1]

Erik Koch
Computational Materials Science
German Research School for Simulation Sciences
Jülich

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

One of the profound *Surprises in Theoretical Physics* [1] is that magnetism is an inherently quantum mechanical effect. Classically, magnetic moments originate from electric currents: A current density $\vec{j}(\vec{r})$ generates a magnetic moment

$$\vec{\mu} = \frac{1}{2}\int \vec{r}\times\vec{j}\,d^3r\,. \tag{1}$$

These moments interact via the dipole-dipole interaction. The magnetostatic interaction energy between two dipoles at a distance R, $\hat{R}$ being the unit-vector from the position of the first to that of the second dipole,

$$\Delta E = \frac{\mu_0}{4\pi}\,\frac{\vec{\mu}_1\cdot\vec{\mu}_2 - 3(\hat{R}\cdot\vec{\mu}_1)(\hat{R}\cdot\vec{\mu_2})}{R^3} = \frac{\vec{\mu}_1\cdot\vec{\mu}_2 - 3(\hat{R}\cdot\vec{\mu}_1)(\hat{R}\cdot\vec{\mu_2})}{4\pi\varepsilon_0 c^2\;R^3} \tag{2}$$

depends on their distance and relative orientation. This can, however, not be the origin of the magnetism found in actual materials: In a classical system charges cannot flow in thermodynamic equilibrium, the celebrated Bohr-van Leeuwen theorem, and hence there are no magnetic moments to begin with [2].
In quantum mechanics, however, non-vanishing charge currents in the ground state are not uncommon: An electron in state $\Psi(\vec{r})$ corresponds to a current density

$$\vec{j}(\vec{r}) = -\frac{e\hbar}{2im_e}\left(\overline{\Psi(\vec{r})}\,\nabla\Psi(\vec{r}) - \Psi(\vec{r})\,\nabla\overline{\Psi(\vec{r})}\right) \tag{3}$$

which, for complex wave function $\Psi(\vec{r})$, is usually non-vanishing. According to (1) it produces a magnetic moment proportional to the expectation value of the angular momentum

$$\vec{\mu}_L = -\frac{e\hbar}{2m_e}\,\langle\vec{L}\,\rangle = -\mu_B\,\langle\vec{L}\,\rangle\;. \tag{4}$$

The constant of proportionality is the Bohr magneton μ_B. In particular, an atomic orbital $|n,l,m\rangle$ has a magnetic moment proportional to the magnetic quantum number $\vec{\mu} = -\mu_B\,m\,\hat{z}$. Also the electron spin $\vec{S}$ carries a magnetic moment

$$\vec{\mu}_S = -g_e\mu_B\,\langle\vec{S}\,\rangle\;. \tag{5}$$

The constant of proportionality between spin and magnetic moment differs from that between orbital momentum and moment by the gyromagnetic ratio g_0. Dirac theory gives $g_e = 2$, which is changed to $g_e \approx 2.0023\ldots$ by QED corrections.
Atomic moments are thus of the order of μ_B. For two such moments at a distance of 1 Å the magnetostatic energy (2) is of the order of 0.05 meV, corresponding to a temperature of less than 1 K. Therefore, magnetic ordering which, e.g., in magnetite (Fe_3O_4), persists till about 860 K, must originate from an interaction other than the magnetostatic interaction of dipoles. Indeed, it is the interplay of electronic properties which are apparently unrelated to magnetism, the Pauli principle in combination with the Coulomb repulsion (Coulomb exchange) as well as the hopping of electrons (kinetic exchange) that leads to an effective coupling between the magnetic moments in a solid.
The basic mechanisms of the exchange coupling are quite simple: Since many-body wave functions must change sign under the permutation of Fermions, electrons of the same spin cannot be at the same position. Electrons of like spin thus tend to avoid each other, i.e., the probability of

finding them close to each other tends to be lower than for electrons of opposite spin (exchange hole). In that sense the Coulomb energy between two electrons depends on their relative spins. By this argument, aligning electron spins tends be energetically favorable. This *Coulomb exchange* is the basis of Hund's first rule. When more than one atom is involved, electrons can hop from one site to the neighbor. This kinetic term is, again, modified by the Pauli principle, as the hopping to an orbital on its neighboring atom will only be possible, if there is not already an electron of the same spin occupying that orbital and by the Coulomb repulsion among the electrons. This is the idea of *kinetic exchange*. When Coulomb exchange and kinetic terms work together we speak of *double exchange*. In that case the electron-hopping serves to mediate the spin-correlation created on an atom to its neighbors.

Exchange mechanisms are idealizations of characteristic situations found in real materials. As such they are approximations, but they afford a simplification of the complicated realistic description, which provides a good basis for thinking about the relevant effects in a real material. We will start by discussing the effect of Coulomb exchange matrix elements (Sec. 1). To keep things simple, we will discuss a two-orbital model and only mention atomic multiplets and Hund's rule, while the full Coulomb vertex is discussed in the lecture of R. Eder in [3]. Next we turn to exchange mechanisms involving also hopping (Sec. 2). We start by looking at the a simple two-site model with two electrons. Focussing on the limit of strong electronic correlations (Coulomb repulsion dominating electron hopping), we introduce the method of downfolding to derive an effective Hamiltonian in which an explicit coupling of the electron spins appears. While conceptually simple, this direct exchange mechanism is rarely found in real materials. There hopping between correlated orbitals is usually mediated by a weakly correlated orbital. This is the superexchange mechanism. The derivation is very similar to that of kinetic exchange. However, the number of states involved, makes explicit book-keeping tedious. To simplify our work, we introduce second quantization as a simple notation of many-electron states. This also enables us to easily discuss double exchange, which combines direct exchange on an atom with coupling to the neighbors via electron hopping. Examples are the superexchange between transition metal atoms bridged by an oxygen at a right angle, which arises from the Coulomb exchange on the oxygen, as well as the exchange in mixed-valence compounds. The competition between kinetic and double exchange is described by the Goodenough-Kanamori rules. For lack of space we will not discuss the mixed-valence double exchange or the exchange mechanisms leading not only to an ordering of spins but to orbital-ordering. These are discussed in my lecture in [3], from which most of the present chapter is drawn.

How exchange gives rise to an effective coupling of momenta is most easily shown for single- or two-site models. To see how these results carry over to solids, we consider the case of direct exchange (Sec. 3). Starting from the Hubbard model we show how taking the limit of strong correlations leads to the t-J-model, which, for half-filling, simplifies to the Heisenberg model. Here we can, of course only discuss the simples of model Hamiltonians describing extended systems. With physical intuition and modern methods it is nowadays possible to construct accurate model Hamiltonians to describe real materials. The state-of-the-art for the construction of such realistic models and their solution is the LDA+DMFT approach. The physical ideas behind this advanced toolbox for treating complex materials and their implementation on modern supercomputers is the subject of the lectures in [4].

We finish by giving a more systematic introduction to the technique of second quantization that formalizes the heuristic approach taken in the preceding sections. This part is taken from my lecture in [5], which discusses the consequences of the indistinguishability of elementary particles in much more detail.

1 Coulomb exchange

The Coulomb repulsion between electrons,

$$H_U = \sum_{i<j} \frac{1}{|\vec{r}_i - \vec{r}_j|}\,, \tag{6}$$

is manifestly spin-independent. Nevertheless, because of the antisymmetry of the many-electron wave function, the eigenenergies of H_U depend on the spin. This is the basis of the multiplet structure in atoms and of Hund's first two rules.
To understand the mechanism of this Coulomb exchange we consider a simple two-electron model. In the spirit of tight-binding (see the lecture of E. Pavarini [6]), we assume that we have solved the two-electron Hamiltonian H_0, replacing the interaction term H_U, e.g., as a self-consistent potential $\sum_i U(\vec{r}_i)$, obtaining an orthonormal set of one-electron eigenstates $\phi_\alpha(\vec{r})$ with eigenvalues ε_α. We now ask for the effect of re-introducing the interaction $H_U - \sum_i U(\vec{r}_i)$. The largest effect we will find for states that are degenerate.
Let us consider two orbitals $\alpha = a,\, b$. Then the two-electron Slater determinants with spins σ and σ'

$$\begin{aligned}\Psi_{a,\sigma;\,b\sigma'}(\vec{r}_1, s_1;\, \vec{r}_2, s_2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_a(\vec{r}_1)\,\sigma(s_1) & \phi_a(\vec{r}_2)\,\sigma(s_2) \\ \phi_b(\vec{r}_1)\,\sigma'(s_1) & \phi_b(\vec{r}_2)\,\sigma'(s_2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} \Big(\phi_a(\vec{r}_1)\,\phi_b(\vec{r}_2)\;\sigma(s_1)\,\sigma'(s_2) - \phi_b(\vec{r}_1)\,\phi_a(\vec{r}_2)\;\sigma'(s_1)\,\sigma(s_2) \Big)\end{aligned} \tag{7}$$

are degenerate eigenstates of H_0 with eigenvalue $\varepsilon_a + \varepsilon_b$, independent of the spin orientations. To see how this degeneracy is lifted, we calculate the matrix elements of H_U in the basis of the Slater determinants $\Psi_{a,\sigma;\,b\sigma'}$.
When both electrons have the same spin ($\sigma = \sigma'$), we can factor out the spin functions

$$\Psi_{a,\sigma;\,b\sigma} = \frac{1}{\sqrt{2}} \Big(\phi_a(\vec{r}_1)\phi_b(\vec{r}_2) - \phi_b(\vec{r}_1)\phi_a(\vec{r}_2) \Big)\; \sigma(s_1)\sigma(s_2) \tag{8}$$

and obtain

$$\left\langle \Psi_{a,\sigma;\,b,\sigma} \left| \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right| \Psi_{a,\sigma;\,b,\sigma} \right\rangle = \frac{1}{2}\,(U_{ab} - J_{ab} - J_{ba} + U_{ba}) = U_{ab} - J_{ab} \tag{9}$$

where the direct terms are the Coulomb integral

$$U_{ab} = \int d^3r_1 \int d^3r_2\, \frac{|\phi_a(\vec{r}_1)|^2\;|\phi_b(\vec{r}_2)|^2}{|\vec{r}_1 - \vec{r}_2|} \tag{10}$$

while the cross terms give the exchange integral

$$J_{ab} = \int d^3r_1 \int d^3r_2\, \frac{\overline{\phi_a(\vec{r}_1)}\,\phi_b(\vec{r}_1)\;\overline{\phi_b(\vec{r}_2)}\,\phi_a(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|}\,. \tag{11}$$

For the states where the electrons have opposite spin ($\sigma' = -\sigma$)

$$\left\langle \Psi_{a,\sigma;\,b,-\sigma} \left| \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right| \Psi_{a,\sigma;\,b,-\sigma} \right\rangle = U_{ab} \tag{12}$$

the diagonal matrix element has no exchange contribution, as the overlap of the spin functions for the cross terms vanish. There are however off-diagonal matrix elements

$$\left\langle \Psi_{a\uparrow;b\downarrow} \left| \frac{1}{|\vec{r}_1 - \vec{r}_2|} \right| \Psi_{a\downarrow;b\uparrow} \right\rangle = -J_{ab} \ . \tag{13}$$

Since H_U does not change the spins, these are the only non-zero matrix elements. In the basis of the states $\Psi_{\uparrow\uparrow}$, $\Psi_{\uparrow\downarrow}$, $\Psi_{\downarrow\uparrow}$ and $\Psi_{\downarrow\downarrow}$ the Coulomb term is thus given by

$$H_U = \begin{pmatrix} U_{ab} - J_{ab} & 0 & 0 & 0 \\ 0 & U_{ab} & -J_{ab} & 0 \\ 0 & -J_{ab} & U_{ab} & 0 \\ 0 & 0 & 0 & U_{ab} - J_{ab} \end{pmatrix} . \tag{14}$$

The triplet states $\Psi_{\uparrow\uparrow}$ and $\Psi_{\downarrow\downarrow}$ are obviously eigenstates of H_U with eigenenergy

$$\Delta\varepsilon_{\text{triplet}} = U_{ab} - J_{ab} \ . \tag{15}$$

Diagonalizing the 2×2 submatrix, we obtain the third triplet state $(\Psi_{\uparrow\downarrow} + \Psi_{\downarrow\uparrow})/\sqrt{2}$ and the singlet state $(\Psi_{\uparrow\downarrow} - \Psi_{\downarrow\uparrow})/\sqrt{2}$

$$\frac{1}{\sqrt{2}} \left(\Psi_{\uparrow\downarrow} - \Psi_{\downarrow\uparrow} \right) = \frac{1}{\sqrt{2}} \Big(\phi_a(\vec{r}_1)\phi_b(\vec{r}_2) + \phi_b(\vec{r}_1)\phi_a(\vec{r}_2) \Big) \ \frac{1}{\sqrt{2}} \Big(|\downarrow\uparrow\rangle - |\uparrow\downarrow\rangle \Big) \tag{16}$$

with energy

$$\Delta\varepsilon_{\text{singlet}} = U_{ab} + J_{ab} \ . \tag{17}$$

To see whether the triplet or the singlet is lower in energy, we need to know the sign of the exchange matrix element. While the Coulomb integral U_{ab}, having a positive integrand, is obviously positive, it is less obvious that also $J_{ab} > 0$. Introducing $\Phi(\vec{r}) = \overline{\phi_a(\vec{r})}\phi_b(\vec{r})$ and rewriting the integral using the convolution theorem as well as the Fourier transform of $1/r$, we obtain [7, 8]:

$$\begin{aligned} J_{ab} &= \int d^3r_1 \, \overline{\Phi(\vec{r}_1)} \underbrace{\int d^3r_2 \, \frac{1}{|\vec{r}_1 - \vec{r}_2|} \, \Phi(\vec{r}_2)}_{= \int dk \, e^{ikr_1} 4\pi/k^2 \Phi(k) \, / (2\pi)^3} && (18) \\ &= \frac{1}{(2\pi)^3} \int d^3k \underbrace{\int d^3r_1 \, e^{i\vec{k}\cdot\vec{r}_1} \overline{\Phi(\vec{r}_1)}}_{= \overline{\Phi(-k)}} \Phi(\vec{k}\,) \, \frac{4\pi}{k^2} && (19) \\ &= \frac{1}{(2\pi)^3} \int d^3k \, |\Phi(\vec{k}\,)|^2 \, \frac{4\pi}{k^2} > 0 && (20) \end{aligned}$$

Thus the triplet states are below the singlet state by an energy $2J_{ab}$. If the ϕ_α are degenerate atomic orbitals, this is an example of Hund's first rule: For an atomic shell, the lowest state will have maximum spin.

Since H_U only contains interactions within the system of electrons, it commutes with the total orbital momentum $[H_U, \vec{L}_{\text{tot}}] = 0$. Obviously it also commutes with the total spin $\vec{S}_{\text{tot}}$. The eigenstates of $H_0 + H_U$ can thus be classified by their quantum numbers L and S. These terms are written as ${}^{2S+1}L$. For p- and d-shells they are listed in table 1. Hund's rules give

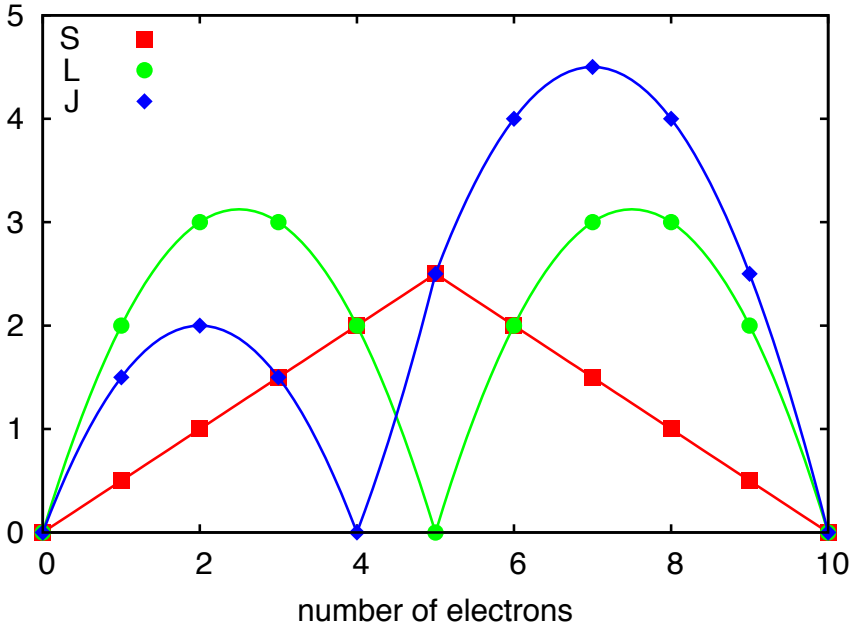


Fig. 1: *Angular momenta of the Hund's rules ground state* $^{2S+1}L_J$ *for* d*-shells.*

the multiplet term with the lowest energy: For a given shell, this lowest state has the largest possible spin (Hund's first rule). If there are several terms of maximum multiplicity, the one with lowest energy has the largest total orbital momentum (Hund's second rule). There is a third Hund's rule, which, however, is not related with the electron-electron repulsion but with spin-orbit coupling: Within L-S coupling H_{SO} splits the atomic orbitals into eigenstates of the total angular momentum $\vec{J} = \vec{L}_{\text{tot}} + \vec{S}_{\text{tot}}$. The multiplets ^{2S+1}L thus split into $^{2S+1}L_J$. The term with the lowest energy is the one with smallest J if the shell is less than half-filled and largest J if it is more than half-filled (Hund's third rule). These rules are illustrated for d-shells in Fig. 1. A more detailed discussion of multiplet effects and the Coulomb interaction in atomic-like systems is given in the lectures [9, 6] and in [10]

s					^{2}**S**									
p^1 or p^5					^{2}**P**									
p^2 or p^4		^1S	^1D						^{3}**P**					
p^3					^2P	^2D					^{4}**S**			
d^1 or d^9					^{2}**D**									
d^2 or d^8		^1S	^1D	^1G					^3P	^{3}**F**				
d^3 or d^7					^2P	$^{2\times}{}^2D$	^2F	^2G	^2H			^4P	^{4}**F**	
d^4 or d^6	$^{2\times}{}^1S$	$^{2\times}{}^1D$	^1F	$^{2\times}{}^1G$	^1I				$^{2\times}{}^3P$	^3D	$^{2\times}{}^3F$	^3G	^3H	^{5}**D**
d^5					^2S	^2P	$^{3\times}{}^2D$	$^{2\times}{}^2F$	$^{2\times}{}^2G$	^2H	^2I	^4P	^4D ^4F ^4G	^{6}**S**

Table 1: *Atomic multiplets for open* s*-,* p*-, and* d*-shells. For terms that appear multiple times the number of distinct terms is indicated. The Hund's rules ground state is indicated in bold.*

2 Kinetic exchange

When electron-hopping plays the main role in the exchange mechanism, we speak of kinetic exchange. In contrast to Coulomb exchange the resulting interactions are usually antiferromagnetic, i.e., they prefer antiparallel spins. The physical principle of kinetic exchange can be understood in a simple two-site system. We discuss this problem in some detail and introduce two key concept along the way: downfolding and second quantization. More realistic exchange mechanisms are then natural generalizations of this simple mechanism [11, 12, 13].

2.1 A toy model

As a toy model, we consider the minimal model of an H_2 molecule. We restrict ourselves to two (orthogonal) orbitals, ϕ_1 and ϕ_2, separated by some distance. If we add an electron to the system, that electron will be able to move between the two orbitals, with a matrix element $-t$. Because we allow the electron to only occupy two s-orbitals, the Hamiltonian is a 2×2 matrix

$$H = \begin{pmatrix} 0 & -t \\ -t & 0 \end{pmatrix} . \tag{21}$$

This tight-binding Hamiltonian is easily diagonalized giving the linear combinations

$$\phi_\pm = \frac{1}{\sqrt{2}} \Big(\phi_1 \pm \phi_2 \Big) \tag{22}$$

as eigenstates with eigenenergies $\varepsilon_\pm = \mp t$. We have written the hopping matrix element as $-t$, so that for $t > 0$ the state without a node, ϕ_+, is the ground state.
Pictorially we can write the basis states by specifying which orbital the electron occupies. For a spin-up electron we then write

$$\phi_1 = |\uparrow , \cdot \rangle \qquad \text{and} \qquad \phi_2 = |\cdot , \uparrow \rangle \tag{23}$$

where we now represent the basis states by where the electron is located.
If there are two electrons in the system, i.e., one electron per orbital, we can again use basis states which just specify, which orbitals the electrons occupy. For two electrons of opposite spin we then find two states where the electrons are in different orbitals

$$|\uparrow , \downarrow\rangle \qquad |\downarrow , \uparrow\rangle \qquad \text{“covalent states”}$$

and two states where the electrons are in the same orbital

$$|\updownarrow , \cdot \rangle \qquad |\cdot , \updownarrow\rangle \qquad \text{“ionic states”.}$$

In this basis the Hamiltonian matrix for our simple model of the H_2 molecule has the form

$$H = \begin{pmatrix} 0 & 0 & -t & -t \\ 0 & 0 & +t & +t \\ -t & +t & U & 0 \\ -t & +t & 0 & U \end{pmatrix} \qquad \begin{matrix} |\uparrow , \downarrow\rangle \\ |\downarrow , \uparrow\rangle \\ |\updownarrow , \cdot \rangle \\ |\cdot , \updownarrow\rangle \end{matrix} \tag{24}$$

As before, moving an electron to a neighboring orbital gives a matrix element $-t$, with an additional sign when the order of the electrons is changed (Fermi statistics!). For the ionic states,

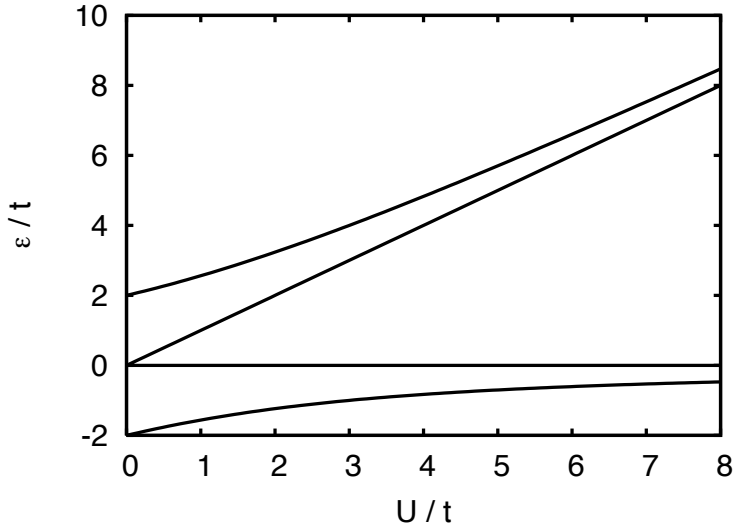


Fig. 2: *Spectrum of the two-site Hubbard model as a function of U. For large U there are two levels with energy close to zero. Their energy difference corresponds to the exchange energy. The remaining two states with ionic character have an energy roughly proportional to U.*

where both electrons are in the same orbital, we have the Coulomb matrix element U. Coulomb matrix elements involving electrons on different sites are, for reasonably large distance between the sites, negligible. So there is no Coulomb exchange, just the local Coulomb repulsion in our model. Diagonalizing H we find the energy spectrum and the corresponding eigenstates:

$$\varepsilon_\pm = \frac{U}{2} \pm \frac{\sqrt{U^2+16\,t^2}}{2} \quad , \quad \Psi_\pm = \frac{\Big(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle - \frac{\varepsilon_\pm}{2t}\big[\,|\updownarrow,\cdot\,\rangle + |\,\cdot\,,\updownarrow\rangle\,\big]\Big)}{\sqrt{2+\varepsilon_\pm^2/(2t^2)}}$$

$$\varepsilon_{\rm cov} = 0 \quad , \quad \Psi_{\rm cov} = \frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle\Big)$$

$$\varepsilon_{\rm ion} = U \quad , \quad \Psi_{\rm ion} = \frac{1}{\sqrt{2}}\Big(|\updownarrow,\cdot\,\rangle - |\,\cdot\,,\updownarrow\rangle\Big)$$

The eigenenergies as a function of U are shown in figure 2.

2.2 Direct exchange

Again, we have found that the energy of two-electron states depends on the relative spin of the electrons. To understand this more clearly we analyze the limit when U is much larger than t. From Fig. 2 we see that there are two states with energies that increase with U. They are the states $\Psi_{\rm ion}$ and Ψ_+ that have considerable contributions of the ionic states. Then there are two states whose energy is close to zero. They are the states that have mainly covalent character. To find the energy and the character of these levels in the limit $U \to \infty$ we can just expand $\varepsilon_- \to -4t^2/U$ and $\varepsilon_+ \to U + 4t^2/U$. We thus see that while the purely covalent state, the spin-triplet state $\Psi_{\rm cov}$, is independent of U, Ψ_- has a slightly lower energy due to some small admixture of the ionic states. In the limit $U \to \infty$ it becomes the maximally entangled state $(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle)/\sqrt{2}$. We see that for large U, Ψ_- cannot be expressed, even approximately, as a Slater determinant, see also Sec. 2.4. This is the reason why strongly correlated systems are so difficult to describe.

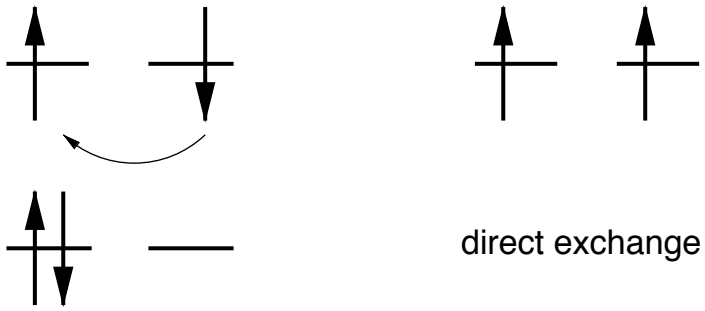


Fig. 3: *Simple picture of direct exchange: The antiparallel alignment of the spins (left) is favored, since it allows the electrons to hop to the neighboring site. For parallel spins (right) hopping is suppressed by the Pauli principle.*

A more instructive method to analyze the large-U limit, which can readily be generalized to more complex situations, where we can no longer diagonalize the full Hamiltonian, is the downfolding technique. The mathematical background is explained in the appendix. The idea of downfolding is to partition the Hilbert space into parts that are of interest, here the low-energy convalent type states, and states that should be projected out, here the high-energy ionic states. With this partitioning we can view the Hamitonian matrix (24) as built of 2×2 submatrices. Calculating the inverse on the space of covalent states (see Eqn. (90) in the appendix) we find an *effective Hamiltonian* which now operates on the covalent states only:

$$H_{\text{eff}}(\varepsilon) = \begin{pmatrix} -t & -t \\ +t & +t \end{pmatrix} \begin{pmatrix} \varepsilon - U & 0 \\ 0 & \varepsilon - U \end{pmatrix}^{-1} \begin{pmatrix} -t & +t \\ -t & +t \end{pmatrix} \approx -\frac{2t^2}{U} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \qquad (25)$$

In the last step we have made an approximation by setting ε to zero, which is roughly the energy of the states with covalent character.
The process of eliminating the ionic states thus gives rise to an effective interaction between the covalent states, which was not present in the original Hamiltonian (24). Diagonalizing the effective Hamiltonian, we find

$$\varepsilon_s = -\frac{4t^2}{U} \quad , \qquad \Psi_s = \frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle\Big)$$
$$\varepsilon_t = 0 \qquad , \qquad \Psi_t = \frac{1}{\sqrt{2}}\Big(|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle\Big)$$

These states correspond to the singlet and triplet states in the hydrogen molecule. Here the singlet-triplett splitting is $2J_{\text{direct}} = -4t^2/U$. The other states in the triplet are those with two electrons of parallel spin: $|\uparrow,\uparrow\rangle$ and $|\downarrow,\downarrow\rangle$. They, of course, also have energy zero, as hopping is impossible due to the Pauli principle.
To understand the nature of the effective interaction in the low-energy Hamiltonian we observe that the off-diagonal matrix elements in (25) correspond to flipping the spin of both electrons ("exchange"). Remembering that

$$\vec{S}_1 \cdot \vec{S}_2 = S_1^z S_2^z + \frac{1}{2}\Big(S_1^+ S_2^- + S_1^- S_2^+\Big) \qquad (26)$$

we see that the effective interaction will contain a spin-spin coupling term.

2.3 Second quantization for pedestrians

A more systematic way for obtaining the form of the effective interaction is by using second quantization, which will also help us simplify our notation. In second quantization we use operators to specify in which orbital an electron is located. As an example, $c^\dagger_{1,\uparrow}$ puts a spin-up electron in orbital ϕ_1. Denoting the system with no electrons by $|0\rangle$, the basis states that we have considered so far are written as

$$\begin{aligned} |\uparrow,\cdot\rangle &= c^\dagger_{1\uparrow}|0\rangle \\ |\cdot,\uparrow\rangle &= c^\dagger_{2\uparrow}|0\rangle \end{aligned}$$

for the single-electron states, and

$$\begin{aligned} |\uparrow,\downarrow\rangle &= c^\dagger_{2\downarrow}c^\dagger_{1\uparrow}|0\rangle \\ |\downarrow,\uparrow\rangle &= c^\dagger_{2\uparrow}c^\dagger_{1\downarrow}|0\rangle \\ |\updownarrow,\cdot\rangle &= c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}|0\rangle \\ |\cdot,\updownarrow\rangle &= c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}|0\rangle \end{aligned} \tag{27}$$

for the two-electron states. In order to describe the hopping of an electron from one orbital to another, we introduce operators that annihilate an electron. For example $c_{1\uparrow}$ removes a spin-up electron from orbital ϕ_1. The hopping of an up electron from ϕ_1 to ϕ_2 is thus described by the operator $c^\dagger_{2\uparrow}c_{1\uparrow}$ that first takes an electron out of orbital 1 and then creates one in orbital 2. The Hamiltonian for a spin-up electron hopping between two orbitals can thus be written as

$$H = -t\Big(c^\dagger_{1\uparrow}c_{2\uparrow} + c^\dagger_{2\uparrow}c_{1\uparrow}\Big)\,. \tag{28}$$

Calculating the matrix elements with the single-electron basis states, we recover the matrix (21). For the calculation we need to know that the operators that describe the electrons *anticommute*. This reflects the fact that a many-electron wave function changes sign when two electrons are exchanged. Using the notation $\{a,b\} = ab + ba$ we have

$$\Big\{c_{i\sigma},\, c_{j\sigma'}\Big\} = 0 \qquad \Big\{c^\dagger_{i\sigma},\, c^\dagger_{j\sigma'}\Big\} = 0 \qquad \Big\{c_{i\sigma},\, c^\dagger_{j\sigma'}\Big\} = \delta_{i,j}\delta_{\sigma,\sigma'}$$

Moreover, trying to annihilate an electron in a state where there is no electron, results in zero: $c_{i\sigma}|0\rangle = 0$. Finally, as the notation implies, $c^\dagger_{i\sigma}$ is the adjoint of $c_{i\sigma}$ and $\langle 0|0\rangle = 1$.
To describe the Coulomb repulsion between two electrons in the same orbital we use that $n_{i\sigma} = c^\dagger_{i\sigma}c_{i\sigma}$ returns 0 when operating on a basis state with no spin-σ electron in orbital ϕ_i, and has eigenvalue 1 for a basis state with a spin-σ electron in orbital ϕ_i. It is thus called the occupation-number operator. The Coulomb repulsion in orbital ϕ_1 is then described by the operator $Un_{1\uparrow}n_{1\downarrow}$, which is non-zero only when there is a spin-up *and* a spin-down electron in ϕ_1. The Hamiltonian for our two-orbital model, where both up- and down-spin electrons can hop, and including the Coulomb repulsion for two electrons in the same orbital, is thus given by

$$\begin{aligned} H &= -t\left(c^\dagger_{1\uparrow}c_{2\uparrow} + c^\dagger_{2\uparrow}c_{1\uparrow} + c^\dagger_{1\downarrow}c_{2\downarrow} + c^\dagger_{2\downarrow}c_{1\downarrow}\right) + U\Big(n_{1\uparrow}n_{1\downarrow} + n_{2\uparrow}n_{2\downarrow}\Big) \\ &= -t\sum_{i,j,\sigma} c^\dagger_{j\sigma}c_{i\sigma} + U\sum_i n_{i\uparrow}n_{i\downarrow}\,. \end{aligned} \tag{29}$$

You should convince yourself that when you calculate the matrix elements for the two-electron states, you recover the matrix (24). The great advantage of writing the Hamiltonian in second-quantized form is that it is valid for any number of electrons, while the matrix form is restricted to a particular number of electrons.
Coming back to the effective Hamiltonian (25), we can rewrite H_{eff} in second quantized form:

$$\begin{aligned} H_{\text{eff}} &= -\frac{2t^2}{U}\left(c^\dagger_{2\uparrow}c^\dagger_{1\downarrow}c_{1\downarrow}c_{2\uparrow} - c^\dagger_{2\downarrow}c^\dagger_{1\uparrow}c_{1\downarrow}c_{2\uparrow} - c^\dagger_{2\uparrow}c^\dagger_{1\downarrow}c_{1\uparrow}c_{2\downarrow} + c^\dagger_{2\downarrow}c^\dagger_{1\uparrow}c_{1\uparrow}c_{2\downarrow}\right) \\ &= -\frac{2t^2}{U}\left(c^\dagger_{1\downarrow}c_{1\downarrow}c^\dagger_{2\uparrow}c_{2\uparrow} - c^\dagger_{1\uparrow}c_{1\downarrow}c^\dagger_{2\downarrow}c_{2\uparrow} - c^\dagger_{1\downarrow}c_{1\uparrow}c^\dagger_{2\uparrow}c_{2\downarrow} + c^\dagger_{1\uparrow}c_{1\uparrow}c^\dagger_{2\downarrow}c_{2\downarrow}\right) \end{aligned} \tag{30}$$

Looking at equation (95) in the appendix we see that the spin operators are given in second quantization by

$$S^x_i = \frac{1}{2}\left(c^\dagger_{i\uparrow}c_{i\downarrow} + c^\dagger_{i\downarrow}c_{i\uparrow}\right) \qquad S^y_i = -\frac{i}{2}\left(c^\dagger_{i\uparrow}c_{i\downarrow} - c^\dagger_{i\downarrow}c_{i\uparrow}\right) \qquad S^z_i = \frac{1}{2}\left(n_{i\uparrow} - n_{i\downarrow}\right). \tag{31}$$

From this we find (after some calculation) that the effective Hamiltonian can be written in terms of the spin operators

$$H_{\text{eff}} = \frac{4t^2}{U}\left(\vec{S}_1 \cdot \vec{S}_2 - \frac{n_1\, n_2}{4}\right). \tag{32}$$

To conclude, we again find that the completely spin-independent Hamiltonian (29), in the limit of large U, gives rise to a spin-spin interaction. Since the exchange coupling $J = 4t^2/U$ is positive, states with antiparallel spins have lower energy. Thus direct exchange leads to *antiferromagnetism*.
It is important to realize that the singlet-triplet splitting for the effective Hamiltonian really arises from the admixture of ionic states into the singlet. By downfolding we eliminate the high-energy ionic states, i.e., charge fluctuations, from our Hilbert space. The eliminated states then give rise to an effective spin-spin interaction on the new reduced low-energy Hilbert space. We must therefore keep in mind that, when working with the effective Hamiltonian (32), we are considering slightly different states than when working with the original Hamiltonian (29).

2.4 Mean-field treatment

To conclude our discussion of the simplest kinetic exchange mechanism, it is instructive to consider the results of a mean-field treatment. For the two-electron Hamiltonian (24) it is straightforward to find the Hartree-Fock solution by directly minimizing the energy expectation value for a two-electron Slater determinant. The most general ansatz is a Slater determinant constructed from an orbital $\phi(\theta_\uparrow) = \sin(\theta_\uparrow)\,\phi_1 + \cos(\theta_\uparrow)\,\phi_2$ for the spin-up, and $\phi(\theta_\downarrow) = \sin(\theta_\downarrow)\,\phi_1 + \cos(\theta_\downarrow)\,\phi_2$ for the spin-down electron:

$$|\Psi(\theta_\uparrow,\theta_\downarrow)\rangle = \left(\sin(\theta_\downarrow)\,c^\dagger_{1\downarrow} + \cos(\theta_\downarrow)\,c^\dagger_{2\downarrow}\right)\,\left(\sin(\theta_\uparrow)\,c^\dagger_{1\uparrow} + \cos(\theta_\uparrow)\,c^\dagger_{2\uparrow}\right)\,|0\rangle\,. \tag{33}$$

Translating the second quantized states via (27) into the basis used for writing the Hamiltonian matrix (24), we find the expectation value

$$\begin{aligned} \langle\Psi(\theta_\uparrow,\theta_\downarrow)|H|\Psi(\theta_\uparrow,\theta_\downarrow)\rangle &= -2t\,(\sin\theta_\uparrow\sin\theta_\downarrow + \cos\theta_\uparrow\cos\theta_\downarrow)\,(\cos\theta_\uparrow\sin\theta_\downarrow + \sin\theta_\uparrow\cos\theta_\downarrow) \\ &\quad + U\left(\sin^2\theta_\uparrow\sin^2\theta_\downarrow + \cos^2\theta_\uparrow\cos^2\theta_\downarrow\right)\,. \end{aligned} \tag{34}$$

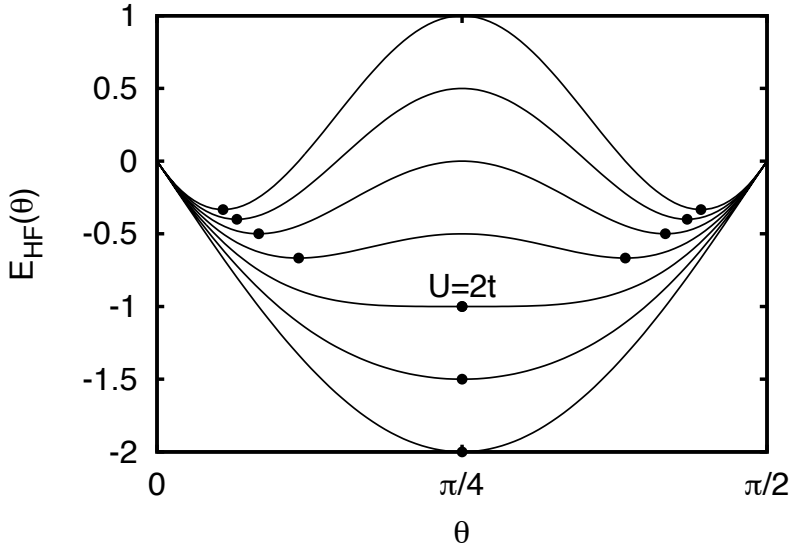


Fig. 4: *Energy expectation value for a Slater determinant $\Psi(\theta, \pi/2-\theta)$ for $U=0,\ t,\ 2t,\ \ldots,\ 6t$. When $U \leq 2t$ the minimum is at $\theta = \pi/4$. This is the Hartree-Fock solution with the bonding orbitals ϕ_+ occupied. For $U \geq 2t$, $\theta = \pi/4$ is still an extremal point (restricted Hartree-Fock solution), but an energy minimum is only attained when the symmetry is broken (unrestricted Hartree-Fock solution).*

If the Slater determinant respects the symmetry of the molecule under the exchange of sites (mirror symmetry of the H_2 molecule), it follows that the Hartree-Fock orbitals for both spins are the bonding state ϕ_+ ($\theta = \pi/4$). This is the *restricted Hartree-Fock* solution. The corresponding energy is $E(\pi/4, \pi/4) = -2t + U/2$. The excited states are obtained by replacing occupied orbitals ϕ_+ with ϕ_-. Alltogether we obtain the restricted Hartree-Fock spectrum

$$\begin{aligned} E(\ \ \pi/4,\ \ \pi/4) &= -2t + U/2 \\ E(\ \ \pi/4, -\pi/4) &= \qquad U/2 \\ E(-\pi/4,\ \ \pi/4) &= \qquad U/2 \\ E(-\pi/4, -\pi/4) &= \ \ 2t + U/2 \end{aligned} \tag{35}$$

Comparing to the energy for a state with both electrons of the same spin ($E = 0$), we see that there is no spin-triplet, i.e., Hartree-Fock breaks the spin symmetry. The states (35) are *spin-contaminated* [14]. Even worse, the Hartree-Fock ground state, and consequently all the states, are independent of U. The weight of the ionic states is always $1/2$, leading to an increase of the energy with $U/2$.

To avoid this, we can allow the Hartree-Fock solution to break the symmetry of the molecule (*unrestricted Hartree-Fock*), putting, e.g., more of the up-spin electron in the orbital on site 1 and more of the down-spin electron in orbital 2. For $U < 2t$ this does not lead to a state of lower energy. For larger U there is a symmetry-broken ground state

$$\Psi_{UHF} = \Psi(\theta, \pi/2 - \theta) \ \text{ with } \quad \theta(U) = \frac{\pi}{4} \pm \frac{1}{2} \arccos\left(\frac{2t}{U}\right) . \tag{36}$$

Its energy is $E_{UHF} = -2t^2/U$. This looks similar to the singlet energy ε_s, however, with a different prefactor. Still there is no triplet state (spin contamination) and, for $U \to \infty$, the overlap with the true singlet ground state goes to $|\langle \Psi_{UHF} | \Psi_- \rangle|^2 = 1/2$. In an extended system the breaking of the symmetry implies long-range order.

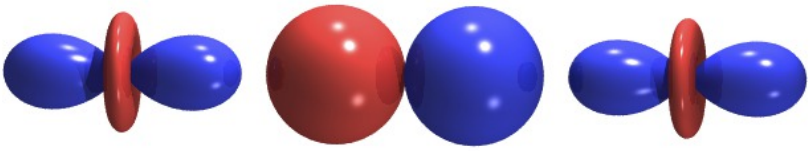


Fig. 5: *In superexchange an oxygen p-orbital mediates the exchange interaction between two transition-metal d-orbitals.*

2.5 Superexchange

For the direct exchange mechanism discussed above, it is crucial that there is hopping between the orbitals. These orbitals are typically localized d-orbitals of transition-metals. However, direct exchange cannot explain the antiferromagnetism of most transition-metal compounds: Since the d-orbitals are so localized, hopping can only occur between orbitals on different atoms that are very close to each other. But most antiferromagnetic insulators are transition-metal *oxides*, so that the transition-metal cations are separated by large oxygen anions. In such a situation, shown in figure 5, direct hopping between the d-orbitals is very unlikely. The concept of direct exchange can, however, be extended to these cases by taking into account hopping via the intermediate p-orbital. This mechanism is called superexchange.
To understand superexchange, we consider two d-orbitals with an oxygen p-orbital in between. We introduce the operator $c^\dagger_{i\sigma}$, which creates a spin-σ electron in the d-orbital at site i, where $i = 1$ denotes the d-orbital on the left and $i = 2$ the one on the right (see figure 5). Likewise $c^\dagger_{p\sigma}$ creates an electron in the p-orbital. The energy of an electron in a d- or p-orbital is ε_d and ε_p, respectively. The Coulomb repulsion between two electrons in a d-orbital is U_d, while we neglect the repulsion between electrons in the p-orbital. Finally, $-t_{pd}$ is the hopping between p and d orbitals. The Hamiltonian for the system of figure 5 is then given by

$$H = \sum_\sigma \left(\varepsilon_d \sum_i n_{i\sigma} + \varepsilon_p \, n_{p\sigma} - t_{pd} \sum_i \left(c^\dagger_{i\sigma} c_{p\sigma} + c^\dagger_{p\sigma} c_{i\sigma} \right) \right) + U_d \sum_i n_{i\uparrow} n_{i\downarrow} \, . \tag{37}$$

In the absence of hopping, the ground state will have singly occupied d-orbitals, corresponding to a positively charged transition-metal ion, and a doubly occupied p-orbital, corresponding to an O^{2-} ion. To study a possible coupling between the spins on the d-orbitals, we first look at the case where both d-spins point upwards (see the far right of Fig. 6). The Hamiltonian matrix in the corresponding Hilbert space is then given by

$$H = \left(\begin{array}{c|cc} 0 & t_{pd} & t_{pd} \\ \hline t_{pd} & U_d + \Delta_{pd} & 0 \\ t_{pd} & 0 & U_d + \Delta_{pd} \end{array} \right) \begin{array}{l} c^\dagger_{2\uparrow} c^\dagger_{p\downarrow} c^\dagger_{p\uparrow} c^\dagger_{1\uparrow} |0\rangle \\ c^\dagger_{2\uparrow} c^\dagger_{p\uparrow} c^\dagger_{1\downarrow} c^\dagger_{1\uparrow} |0\rangle \\ c^\dagger_{2\downarrow} c^\dagger_{2\uparrow} c^\dagger_{p\uparrow} c^\dagger_{1\uparrow} |0\rangle \end{array} \tag{38}$$

where we have chosen $2(\varepsilon_p + \varepsilon_d)$ as the zero of our energy scale and defined $\Delta_{pd} = \varepsilon_d - \varepsilon_p$. The basis states of the Hilbert space are given on the right and the lines indicate the partitioning of the Hilbert space for downfolding. The effective Hamiltonian for parallel spins on d-orbitals is then

$$H_{\text{eff}} = (t_{pd}, t_{pd}) \begin{pmatrix} \varepsilon - (U_d + \Delta_{pd}) & 0 \\ 0 & \varepsilon - (U_d + \Delta_{pd}) \end{pmatrix} \begin{pmatrix} t_{pd} \\ t_{pd} \end{pmatrix} \approx - \frac{2 t_{pd}^2}{U_d + \Delta_{pd}} \tag{39}$$

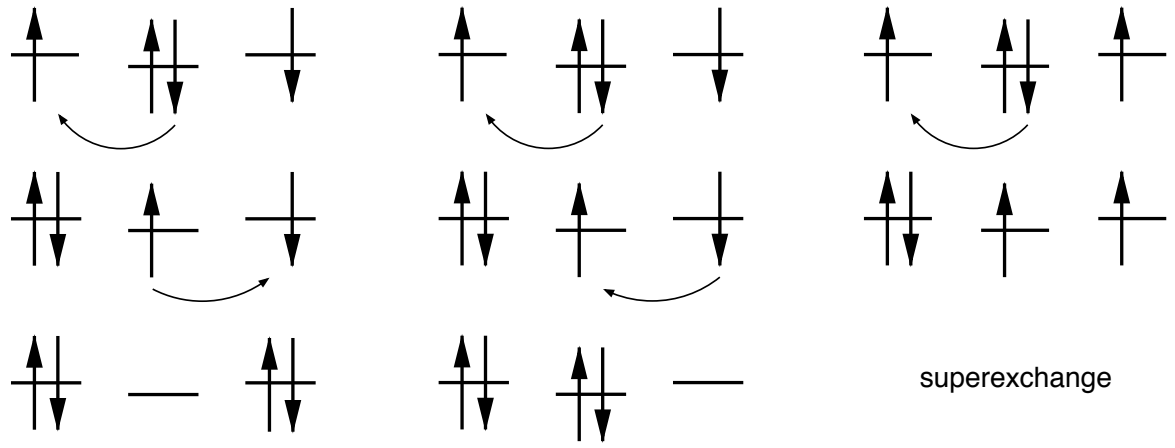


Fig. 6: *Simple picture of superexchange. Here the orbital on the central site is different from the orbitals on the sides. Typically, in the center there is a oxygen p-orbital coupling two d-orbitals. This situation is illustrated in Fig. 5. For antiparallel spins on the d-orbitals there are two ways that two consecutive hopping processes are possible. For parallel spins the Pauli principle suppresses the second hopping process.*

where in the last step we have set ε to zero.
For antiparallel spins the Hilbert space is nine-dimensional. We sort the basis states into groups that are connected by the hopping of one electron. Starting from the two states with singly occupied d-orbitals, the second group has one of the p-electrons transfered to a d-orbital, leading to one doubly occupied d, while the last group has a second electron hopped, leading to either an empty p- or an empty d-orbital. The corresponding Hamiltonian matrix is

$$\left(\begin{array}{cc|cccc|ccc}
0 & 0 & +t_{pd} & +t_{pd} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & +t_{pd} & +t_{pd} & 0 & 0 & 0 \\
\hline
+t_{pd} & 0 & U_d+\Delta_{pd} & 0 & 0 & 0 & -t_{pd} & 0 & -t_{pd} \\
+t_{pd} & 0 & 0 & U_d+\Delta_{pd} & 0 & 0 & 0 & -t_{pd} & -t_{pd} \\
0 & +t_{pd} & 0 & 0 & U_d+\Delta_{pd} & 0 & +t_{pd} & 0 & +t_{pd} \\
0 & +t_{pd} & 0 & 0 & 0 & U_d+\Delta_{pd} & 0 & +t_{pd} & +t_{pd} \\
\hline
0 & 0 & -t_{pd} & 0 & +t_{pd} & 0 & U_d & 0 & 0 \\
0 & 0 & 0 & -t_{pd} & 0 & +t_{pd} & 0 & U_d & 0 \\
0 & 0 & -t_{pd} & -t_{pd} & +t_{pd} & +t_{pd} & 0 & 0 & 2(U_d+\Delta_{pd})
\end{array}\right)
\begin{array}{l}
c^\dagger_{2\downarrow}c^\dagger_{p\downarrow}c^\dagger_{p\uparrow}c^\dagger_{1\uparrow}|0\rangle \\
c^\dagger_{2\uparrow}c^\dagger_{p\downarrow}c^\dagger_{p\uparrow}c^\dagger_{1\downarrow}|0\rangle \\
c^\dagger_{2\downarrow}c^\dagger_{p\uparrow}c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}|0\rangle \\
c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}c^\dagger_{p\downarrow}c^\dagger_{1\uparrow}|0\rangle \\
c^\dagger_{2\uparrow}c^\dagger_{p\downarrow}c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}|0\rangle \\
c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}c^\dagger_{p\uparrow}c^\dagger_{1\downarrow}|0\rangle \\
c^\dagger_{p\downarrow}c^\dagger_{p\uparrow}c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}|0\rangle \\
c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}c^\dagger_{p\downarrow}c^\dagger_{p\uparrow}|0\rangle \\
c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}|0\rangle
\end{array}$$

Downfolding the high energy states with at least one doubly occupied d-orbital, setting $\varepsilon = 0$ and expanding in $1/U_d$ (remembering $(A+\Delta)^{-1} \approx A^{-1}(1-\Delta A^{-1})$), which is equivalent to second-order perturbation theory, leads to

$$\begin{aligned}
H_{\text{eff}} &= H_{00} + T_{01}\left(\varepsilon - \left(H_{11} + T_{12}\left(\varepsilon - H_{22}\right)^{-1}T_{21}\right)\right)^{-1}T_{10} \\
&\approx H_{00} - T_{01}H_{11}^{-1}T_{10} - T_{01}H_{11}^{-1}T_{12}H_{22}^{-1}T_{21}H_{11}^{-1}T_{10} \qquad (40) \\
&= -\frac{2t_{pd}^2}{U_d+\Delta_{pd}}\begin{pmatrix}1&0\\0&1\end{pmatrix} - \frac{2t_{pd}^4}{(U_d+\Delta_{pd})^2}\left(\frac{1}{U_d}+\frac{1}{U_d+\Delta_{pd}}\right)\begin{pmatrix}1&-1\\-1&1\end{pmatrix}. \qquad (41)
\end{aligned}$$

The first term is the same as for parallel spins (39). The additional term is of the same type as that found for the direct exchange mechanism. Again, it can be written in terms of spin operators. In the present case they are the spin operators for the d-orbitals, while the p-orbital does no longer appear in the spin Hamiltonian. The spin coupling is now given by

$$J = \frac{4t_{pd}^4}{(U_d + \Delta_{pd})^2} \left(\frac{1}{U_d} + \frac{1}{U_d + \Delta_{pd}} \right), \tag{42}$$

which reflects that the superexchange mechanism involves four hopping processes (see Fig. 6), while direct exchange only involves two hoppings (see Fig. 3). The hopping process involving only a single doubly occupied d-orbital (middle of Fig. 6) is a generalization of the simple direct exchange with an effective hopping $t_{\rm eff} = t_{pd}^2/(U_d + \Delta_{pd})$ between the d-orbitals and gives the first term, $4t_{\rm eff}^2/U_d$, in (42), while the hopping process involving two occupied d-orbitals (left in Fig. 6) gives the second term $4t_{pd}^4/(U_d + \Delta_{pd})^3$.

2.6 Ferromagnetic superexchange

In the discussion of superexchange we have, so far, assumed that the oxygen ion lies between the two d-orbitals. This 180^o geometry is shown on the left of Fig. 7. The situation is quite different, when the oxygen forms a 90^o bridge between the two d-orbitals, see the right of Fig. 7. By symmetry, there is only hopping between the d- and the p-orbital that point towards each other (see, e.g., the discussion of the Slater-Koster integrals in the lecture of E. Pavarini in [3]). As there is also no hopping between the p-orbitals on the same site, the Hamiltonian for the system separates into two parts, one involving only the d orbital on site 1 and the p_x orbital and the other only involving d on site 2 and p_y, e.g.:

$$H_1 = \begin{pmatrix} 0 & +t_{pd} \\ +t_{pd} & U_d + \Delta_{pd} \end{pmatrix} \quad \begin{matrix} c_{x\downarrow}^\dagger c_{x\uparrow}^\dagger c_{1\downarrow}^\dagger |0\rangle \\ c_{x\downarrow}^\dagger c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger |0\rangle \end{matrix} \tag{43}$$

Since it is not possible for an electron on site 1 to reach site 2, none of the superexchange processes discussed above are operational. Nevertheless, the energy for the system depends on the relative orientation of the electron spins in the two d-orbitals. To see this, we have to remember that Coulomb exchange prefers a triplet for two electrons in different orbitals on the same site (Hund's first rule). Including J_{xy} on the oxygen (but neglecting U_p for simplicity), we get, for the triplet state with two up-electrons, the Hamiltonian

$$\left(\begin{array}{c|cc|c} 0 & t_{pd} & t_{pd} & 0 \\ \hline t_{pd} & U_d + \Delta_{pd} & 0 & t_{pd} \\ t_{pd} & 0 & U_d + \Delta_{pd} & t_{pd} \\ \hline 0 & t_{pd} & t_{pd} & 2(U_d + \Delta_{pd}) - J_{xy} \end{array} \right) \quad \begin{matrix} c_{1\uparrow}^\dagger c_{x\downarrow}^\dagger c_{x\uparrow}^\dagger c_{y\downarrow}^\dagger c_{y\uparrow}^\dagger c_{2\uparrow}^\dagger |0\rangle \\ c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger c_{x\uparrow}^\dagger c_{y\downarrow}^\dagger c_{y\uparrow}^\dagger c_{2\uparrow}^\dagger |0\rangle \\ c_{1\uparrow}^\dagger c_{x\downarrow}^\dagger c_{x\uparrow}^\dagger c_{y\uparrow}^\dagger c_{2\downarrow}^\dagger c_{2\uparrow}^\dagger |0\rangle \\ c_{1\downarrow}^\dagger c_{1\uparrow}^\dagger c_{x\uparrow}^\dagger c_{y\uparrow}^\dagger c_{2\downarrow}^\dagger c_{2\downarrow}^\dagger |0\rangle \end{matrix} . \tag{44}$$

The first state has the two up-electrons on the d-orbitals. The second group of states has one d-orbital doubly occupied, while the last state has both d doubly occupied, i.e., two electrons on the two p-orbitals – the situation discussed in Sec. 1. Calculating the effective Hamiltonian as in (40) gives the energy of the triplet state

$$H_{\rm eff} = -\frac{2t_{pd}^2}{U_d + \Delta_{pd}} - \frac{4t_{pd}^4}{(U_d + \Delta_{pd})^2} \frac{1}{2(U_d + \Delta_{pd}) - J_{xy}} . \tag{45}$$

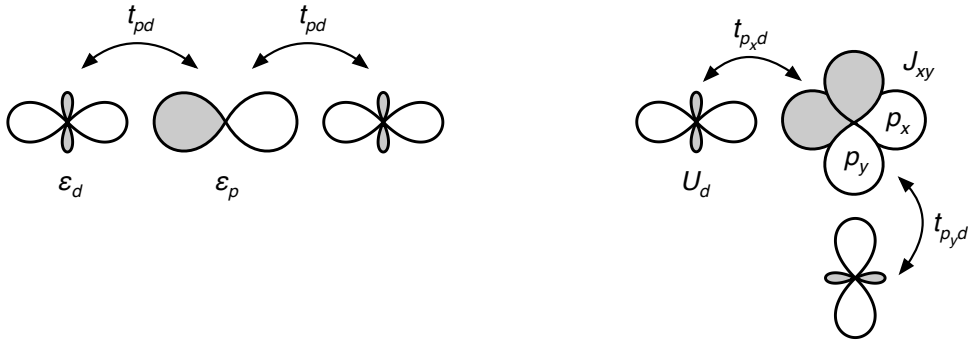


Fig. 7: *Dependence of superexchange on geometry: When the d-orbitals interact via an oxygen in-between (the 180^o geometry shown on the left), both d-orbitals couple to the same p-orbital, while the hopping to the two other p-orbitals vanishes by symmetry. The result is antiferromagnetic superexchange. When the angle of the M-O-M group is 90^o (right), the d-orbitals couple to orthogonal p-orbitals, making it impossible for an electron on one d-orbital to reach the d-orbital on the other site. In this case, superexchange is mediated via the Coulomb exchange on the connecting oxygen.*

Starting from singly occupied d orbitals with opposite spin, we obtain

$$\left(\begin{array}{cc|cccc|cc}
0 & 0 & t_{pd} & 0 & t_{pd} & 0 & 0 & 0 \\
0 & 0 & 0 & t_{pd} & 0 & t_{pd} & 0 & 0 \\
\hline
t_{pd} & 0 & U_d+\Delta_{pd} & 0 & 0 & 0 & t_{pd} & 0 \\
0 & t_{pd} & 0 & U_d+\Delta_{pd} & 0 & 0 & 0 & t_{pd} \\
t_{pd} & 0 & 0 & 0 & U_d+\Delta_{pd} & 0 & t_{pd} & 0 \\
0 & t_{pd} & 0 & 0 & 0 & U_d+\Delta_{pd} & 0 & t_{pd} \\
\hline
0 & 0 & t_{pd} & 0 & t_{pd} & 0 & 2(U_d+\Delta_{pd}) & -J_{xy} \\
0 & 0 & 0 & t_{pd} & 0 & t_{pd} & -J_{xy} & 2(U_d+\Delta_{pd})
\end{array}\right)
\begin{array}{l}
c^\dagger_{1\uparrow}c^\dagger_{x\downarrow}c^\dagger_{x\uparrow}c^\dagger_{y\downarrow}c^\dagger_{y\uparrow}c^\dagger_{2\downarrow}|0\rangle \\
c^\dagger_{1\downarrow}c^\dagger_{x\downarrow}c^\dagger_{x\uparrow}c^\dagger_{y\downarrow}c^\dagger_{y\uparrow}c^\dagger_{2\uparrow}|0\rangle \\
c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{x\uparrow}c^\dagger_{y\downarrow}c^\dagger_{y\uparrow}c^\dagger_{2\downarrow}|0\rangle \\
c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{x\downarrow}c^\dagger_{y\downarrow}c^\dagger_{y\uparrow}c^\dagger_{2\uparrow}|0\rangle \\
c^\dagger_{1\uparrow}c^\dagger_{x\downarrow}c^\dagger_{x\uparrow}c^\dagger_{y\downarrow}c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}|0\rangle \\
c^\dagger_{1\downarrow}c^\dagger_{x\downarrow}c^\dagger_{x\uparrow}c^\dagger_{y\uparrow}c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}|0\rangle \\
c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{x\uparrow}c^\dagger_{y\downarrow}c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}|0\rangle \\
c^\dagger_{1\downarrow}c^\dagger_{1\uparrow}c^\dagger_{x\downarrow}c^\dagger_{y\uparrow}c^\dagger_{2\downarrow}c^\dagger_{2\uparrow}|0\rangle
\end{array}$$

giving the effective Hamiltonian

$$H_{\text{eff}} = -\frac{2t_{pd}^2}{U_d+\Delta_{pd}}\begin{pmatrix}1&0\\0&1\end{pmatrix} - \frac{4t_{pd}^4}{(U_d+\Delta_{pd})^2}\frac{1}{4(U_d+\Delta_{pd})^2-J_{xy}^2}\begin{pmatrix}2(U_d+\Delta_{pd}) & +J_{xy}\\ +J_{xy} & 2(U_d+\Delta_{pd})\end{pmatrix}.$$

Rearranging the matrices, we can bring this to the canonical form

$$\begin{aligned} H_{\text{eff}} = &-\left(\frac{2t_{pd}^2}{U_d+\Delta_{pd}} + \frac{4t_{pd}^4}{(U_d+\Delta_{pd})^2}\frac{1}{2(U_d+\Delta_{pd})-J_{xy}}\right) \\ &+ \frac{4t_{pd}^4}{(U_d+\Delta_{pd})^2}\frac{J_{xy}}{4(U_d+\Delta_{pd})^2-J_{xy}^2}\begin{pmatrix}1&-1\\-1&1\end{pmatrix}. \end{aligned} \tag{46}$$

The first term is just the energy of the triplet state (45). The second gives the difference in energy to the singlet. Despite the fact that the electrons cannot be transferred between the d orbitals we thus get a singlet-triplet splitting. This coupling of the spins originates from the states with both d-orbitals doubly occupied: the two remaining electrons, one each on the p_x- and p_y-orbital, respectively, form a triplet of energy $2J_{xy}$ lower than that of the singlet (see Eqn. (15)). When the electrons hop back from the d-orbital, the entanglement of the spins is transferred to the remaining electron on the d. Originating from the Coulomb exchange on the

oxygen, the exchange coupling is ferromagnetic

$$J = -\frac{4t_{pd}^4}{(U_d + \Delta_{pd})^2} \frac{2J_{xy}}{4(U_d + \Delta_{pd})^2 - J_{xy}^2} . \tag{47}$$

It tends to be significantly weaker than the antiferromagnetic 180^o superexchange coupling (42). When the angle of the M-O-M group is larger than 90^o, hopping to both p-orbitals becomes possible according to the Slater-Koster rules and the antiferromagnetic superexchange processes of Fig. 6 start to compete with the ferromagnetic superexchange mediated by the Coulomb exchange on the oxygen. This is one basis of the Goodenough-Kanamori rules [11, 15].

3 Extended systems

3.1 Hubbard model and Mott transition

We now turn to *extended* systems. For this we consider the Hubbard model [16] on an infinite lattice. Note that now the Hilbert space is infinitely dimensional, so we can no longer write down the Hamiltonian in its matrix form but have to rely on the second quantized form (29)

$$H = -t \sum_{i,j,\sigma} c_{j\sigma}^\dagger c_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} . \tag{48}$$

As in our toy model we still assume that each atom has only a single relevant orbital. There are links between the neighboring atoms with matrix elements t, which can be intuitively interpreted as hopping from site to site. In the absence of other terms the hopping gives rise to a band. A second energy scale is given by the Coulomb repulsion U between two electrons on the same atom. If this on-site Coulomb repulsion is comparable to or even larger than the band width, the electrons can no longer be considered independent; since the double occupation of an atom is energetically very costly, the movement of an electron will be hindered by the Coulomb repulsion. One says that the electrons move in a *correlated* way. We should note that also the Pauli principle hinders the movement of an electron. This effect can, however, be efficiently described by constructing a Slater determinant of independent-electron wave functions. Correlations, on the other hand, are notoriously difficult to describe since no simple wave functions for such systems are available. In the case of strong correlations, i.e., for $U \gg t$, we will treat the hopping as a perturbation. This is called the *atomic limit*, since the sites are almost independent atoms. Thus it is most appropriate to describe strongly correlated electrons in a local picture, i.e., in terms of electron configurations, which are the states that diagonalize the Coulomb term.

The physics described by the Hubbard model is the interplay between kinetic energy and Coulomb repulsion. Writing the Hubbard-Hamiltonian either in real or in k-space

$$\begin{aligned} H = & -t \sum_{i,j,\sigma} c_{j\sigma}^\dagger c_{i\sigma} + U \sum_i n_{i\uparrow} n_{i\downarrow} \\ = & \sum_{k\sigma} \varepsilon_k \, c_{k\sigma}^\dagger c_{k\sigma} + \frac{U}{M} \sum_{k,k',q} c_{k\uparrow}^\dagger c_{k-q\uparrow} c_{k'\downarrow}^\dagger c_{k'+q\downarrow} , \end{aligned}$$

where M is the number of lattice sites, we see that there are obviously two limiting cases: There is the non-interacting- or band-limit, when $t \gg U$. In that case, only the hopping term

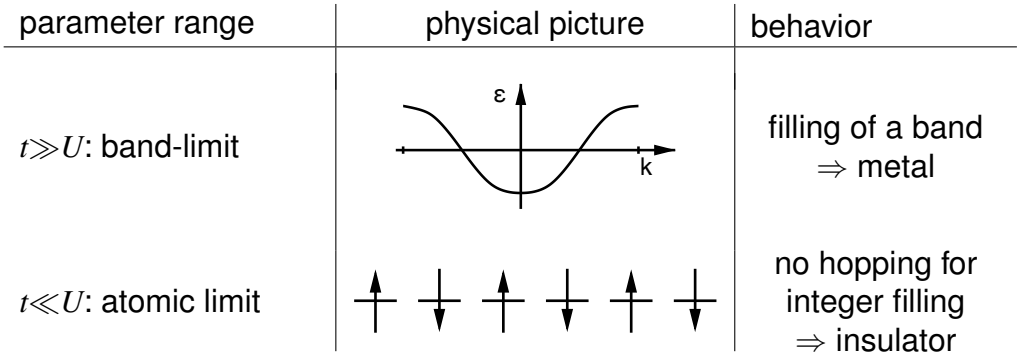


Fig. 8: *Metal-insulator transition for half-filling, i.e., one electron per site.*

survives, i.e., there are no interactions, and the Hamiltonian can be solved easily in k-space. The energy levels then form a band and the system is metallic, except when the band is completely filled. In the opposite case, the atomic limit, the interaction term dominates. In that limit, to minimize the Coulomb energy, the electrons will be distributed over the lattice sites as uniformly as possible. For a non-degenerate, half-filled system this means, that every site carries exactly one electron, and hopping is suppressed, because it would create a doubly occupied site, which would increase the energy by $U \gg t$. Thus in the atomic limit the half-filled system will be an insulator. Clearly, in-between these two limiting cases there must be, at some value U_c, the so-called *critical* U, a transition from a metallic to an insulating state — the Mott transition [17]. Usually this transition is expected when U becomes of the order of the (non-interacting) band width W.
As the criterion for determining the metal-insulator transition we can use the opening of the gap for charge-carrying single-electron excitations

$$E_g = E(N+1) - 2E(N) + E(N-1)\,, \tag{49}$$

where $E(N)$ denotes the total energy of a cluster of M atoms with N electrons. For the half-filled system we have $N = M$. It is instructive to again consider the two limiting cases. In the non-interacting limit the total energy is given by the sum over the eigenvalues of the hopping Hamiltonian $\sum_{n:\mathrm{occ}} \varepsilon_n$. Thus, in the non-interacting limit $E_g^{\mathrm{band}} = \varepsilon_{N+1} - \varepsilon_N$, which, for a partly filled band, will vanish in the limit of infinite system size. On the other hand, in the atomic limit, the Coulomb energy for a single site with n electrons is $Un(n-1)/2$. Thus, for half-filling of we have

$$E_g^{\mathrm{atml}} = U\ , \tag{50}$$

i.e., the insulating state in the atomic limit is characterized by a finite gap.
For an infinite system the gap E_g can be rewritten in terms of the chemical potential. In the thermodynamic limit ($M \to \infty$ with N/M constant) we have to distinguish two types: the energy needed to add an electron to the system (electron affinity)

$$\mu^+ = \lim(E(N+1) - E(N)) = \left.\frac{d\varepsilon(n)}{dn}\right|_{n\searrow 1}\ , \tag{51}$$

and the energy required to extract an electron from the system (ionization energy)

$$\mu^- = \lim(E(N) - E(N-1)) = \left.\frac{d\varepsilon(n)}{dn}\right|_{n\nearrow 1}\ . \tag{52}$$

The gap is then given by the discontinuity in the left- and right-derivative of the energy per site $\varepsilon(n) = \lim E(N)/M$: $E_g = \mu^+ - \mu^-$.

3.2 Heisenberg model

We now consider the Hubbard model in the limit of large U. This is the generalization of the discussion of direct kinetic exchange in Sec. 2.2 to an extended system. For large U we work with the electron configurations, in which the interaction term is diagonal. Configurations with doubly occupied sites will have energies of the order of U or larger, so these are the configurations that we would like to project out. For downfolding we thus partition the configuration basis, and hence the Hilbert space, into the set of low-energy states which have no doubly occupied sites

$$S = \left\{ |n_{1\uparrow}, n_{1\downarrow}, n_{2\uparrow}, n_{2\downarrow}, \ldots\rangle \,\middle|\, \forall i : \, n_{i\uparrow} + n_{i\downarrow} \leq 1 \right\} \tag{53}$$

and the set of high-energy states with one or more doubly occupied sites

$$D = \left\{ |n_{1\uparrow}, n_{1\downarrow}, n_{2\uparrow}, n_{2\downarrow}, \ldots\rangle \,\middle|\, \exists i : \, n_{i\uparrow} + n_{i\downarrow} = 2 \right\} \ . \tag{54}$$

The hopping term T, which for large U is a perturbation to the interaction term I, couples the subspaces by hopping an electron into or out of a doubly occupied site. In addition it lifts the degeneracies within the subspaces. Hence the Hamiltonian can be partitioned as (note that $I \equiv 0$ on subspace S)

$$\hat{H} = \begin{pmatrix} P_S \, T \, P_S & P_S \, T \, P_D \\ P_D \, T \, P_S & P_D \, (T + I) \, P_D \end{pmatrix} \ , \tag{55}$$

Since we are dealing with an extended system, the subspaces are infinite, so we cannot write the Hamiltonian on the subspaces as matrices. Instead we restrict the operators to the appropriate subspace by using projection operators, P_S projecting on the low-energy configurations S, P_D projecting on D. Just like in 2.2 we can then write down an effective Hamiltonian operating on the low-energy configurations only:

$$H_{\text{eff}} = P_S \, T \, P_S + P_S \, T \, P_D \ \left[P_D \left(\varepsilon - (I + T) \right) \, P_D \right]^{-1} P_D \, T \, P_S \ , \tag{56}$$

Unlike in the derivation of direct exchange, for the extended system we have no way of calculating the inverse in the second term explicitly. We can, however, expand in powers of t/U. This is Kato's method for perturbation theory (see, e.g., section 16.3 of [18]). Essentially we only need to consider configurations with a *single* double-occupancy – these correspond to the states of lowest energy in D. On this subspace the interaction term is diagonal with eigenvalue U and can thus be easily inverted. We then obtain the Hamiltonian

$$H_{t-J} = P_S \left[T - \frac{t^2}{U} \sum_{\langle ij \rangle \langle jk \rangle \sigma\sigma'} c^\dagger_{k\sigma'} c_{j\sigma'} \, n_{j\uparrow} n_{j\downarrow} \, c^\dagger_{j\sigma} c_{i\sigma} \right] \ P_S \ , \tag{57}$$

which is called the t-J Hamiltonian. The first term describes the hopping, constrained to configurations with no doubly occupied sites. Thus it essentially describes the hopping of empty sites (holes). To understand what the second term does, we observe that, because of the operators $n_{j\uparrow} n_{j\downarrow}$, there are only contributions for states with a singly occupied site j: $n_{j\sigma} = 0$ and $n_{j,-\sigma} = 1$. After applying the second term, site j will again be singly occupied with $n_{j\sigma'} = 0$ and $n_{j,-\sigma'} = 1$. Hence, for $\sigma \neq \sigma'$ the spin on site j will be flipped. Moreover, we distinguish the contributions where only two different sites are involved ($k = i$) from the three-site terms ($k \neq i$). The terms for $k = i$ are just the ones we already know from the kinetic exchange

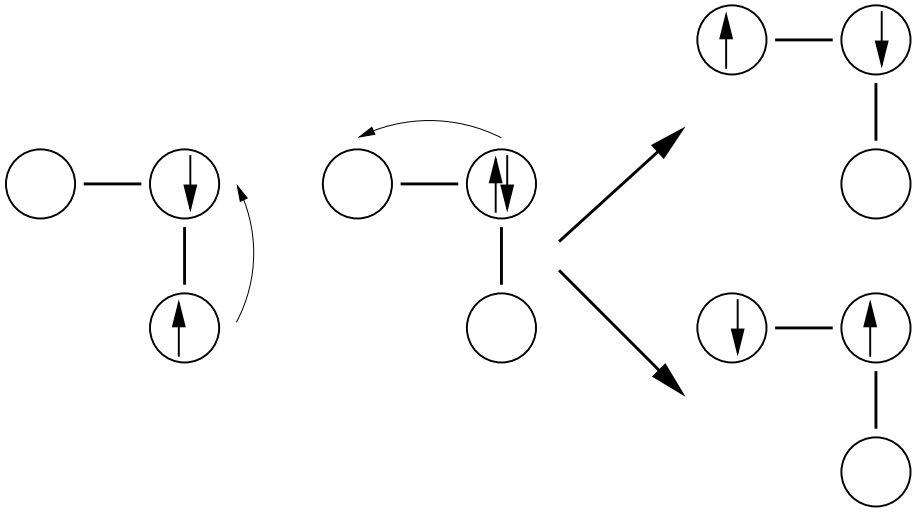


Fig. 9: *Processes contained in the three-site term T': indirect hopping processes to a second-nearest neighbor site with an intermediate (virtual) doubly occupied state. In the first process the two hopping processes are performed by the same electron, in the second process each electron hops once and thus the spin on the intermediate site is flipped.*

mechanism. The three-site terms describe a second-nearest neighbor hopping of an electron from site i to site k via a singly occupied intermediate site j. For $\sigma = \sigma'$ the spin of the hopping electron is opposite to that on the intermediate site. For $\sigma \neq \sigma'$ the spin of the intermediate site is flipped – as is that of the hopping electron. This is shown in Fig. 9. The t-J Hamiltonian is

$$H_{t-J} = P_S \, [T + H_H + T'] \, P_S \tag{58}$$

with

$$T = -t \sum_{\langle ij \rangle, \sigma} c^\dagger_{j\sigma} c_{i\sigma} \tag{59}$$

$$H_H = \frac{4t^2}{U} \sum_{\langle ij \rangle} \left(\vec{S}_j \cdot \vec{S}_i - \frac{n_i n_j}{4} \right) \tag{60}$$

$$T' = -\frac{t^2}{U} \sum_{\substack{\langle ij \rangle \langle jk \rangle \\ i \neq k}} \sum_{\sigma} \left(c^\dagger_{k\sigma} (1 - n_{j\sigma}) c_{i\sigma} - c^\dagger_{k,-\sigma} c^\dagger_{j\sigma} c_{j,-\sigma} c_{i\sigma} \right) n_{j,-\sigma} \tag{61}$$

In the case of half-filling, when $n_i = 1$, all hopping processes are suppressed, i.e., the projection P_S annihilates T and T'. Thus for a Mott insulator the t-J model reduces to the spin $1/2$ Heisenberg model

$$H_H = J \sum_{\langle ij \rangle} \vec{S}_j \cdot \vec{S}_i + \text{const.} \tag{62}$$

with the exchange coupling $J = 4t^2/U$ given by the direct kinetic exchange mechanism. We again stress that the spin-spin interaction is a result of projecting out the states with double occupancies.

4 Second quantization

The first object to be successfully quantized was the electron. It was no longer described as a classical point-particle but by a quantum mechanical Schrödingier field. Later, for studying the interaction of radiation with matter, also the electromagnetic field had to be quantized, giving rise to quantum particles – photons. This process, pioneered by Dirac [19], was called the second quantization. Shortly after, Jordan, Klein, and Wigner used a similar approach to quantize the Schrödingier field and found that it could be used to write antisymmetric states in a very convenient way using particle-type operators [20, 21].
When working with Slater determinants of the form

$$\Phi_{\alpha_1\cdots\alpha_N}(x_1,\ldots,x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\alpha_1}(x_1) & \varphi_{\alpha_2}(x_1) & \cdots & \varphi_{\alpha_N}(x_1) \\ \varphi_{\alpha_1}(x_2) & \varphi_{\alpha_2}(x_2) & \cdots & \varphi_{\alpha_N}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(x_N) & \varphi_{\alpha_2}(x_N) & \cdots & \varphi_{\alpha_N}(x_N) \end{vmatrix} \tag{63}$$

we are working in a real-space basis. Like in fundamental quantum mechanics, it is, however, often useful to abstract from a specific basis and work with abstract states: Instead of a wave function $\varphi_\alpha(x)$, we write a Dirac state $|\alpha\rangle$. Second quantization allows us to do the same for Slater determinants.
Let us consider a Slater determinant for two electrons, one in state $\varphi_\alpha(x)$, the other in state $\varphi_\beta(x)$. It is simply the antisymmetrized product of the two states

$$\Phi_{\alpha\beta}(x_1,x_2) = \frac{1}{\sqrt{2}}\left(\varphi_\alpha(x_1)\varphi_\beta(x_2) - \varphi_\beta(x_1)\varphi_\alpha(x_2)\right). \tag{64}$$

We could do the same for Dirac states, defining a two-particle Dirac state

$$|\alpha,\beta\rangle := \frac{1}{\sqrt{2}}\left(|\alpha\rangle|\beta\rangle - |\beta\rangle|\alpha\rangle\right).$$

The idea of second quantization is then to specify the states using operators

$$c_\beta^\dagger c_\alpha^\dagger|0\rangle = |\alpha,\beta\rangle. \tag{65}$$

When these operators change sign when they are reordered, antisymmetry of the wave function will be automatically ensured

$$|\alpha,\beta\rangle = c_\beta^\dagger c_\alpha^\dagger|0\rangle = -c_\alpha^\dagger c_\beta^\dagger|0\rangle = -|\beta,\alpha\rangle. \tag{66}$$

Naturally, this also implies the Pauli principle for the special case $\beta = \alpha$..

4.1 Creation and annihilation operators

To arrive at the formalism of second quantization we postulate a set of operators that have certain reasonable properties. We then verify that we can use operators with these properties to represent Slater determinants. We start by motivating the properties of the new operators.
To be able to construct many-electron states, we start from the simplest such state: $|0\rangle$ the state with no electron, i.e., the *vacuum state*, which we assume to be normalized $\langle 0|0\rangle = 1$. Next

we introduce for each single-electron state $|\alpha\rangle$ (corresponding to an orbital $\varphi_\alpha(x)$) an operator $c_\alpha^\dagger$. We call it a c*reation operator*, since we ask that applying $c_\alpha^\dagger$ to an N-electron state adds an electron in state $|\alpha\rangle$ to that state, making it an $N+1$ electron state. In effect, the operator should be constructed such as to mimic the effect of adding an extra column φ_α and an extra row x_{N+1} to the Slater determinant (63). Since the order in which we add rows/columns matters for the sign of the Slater determinant, we postulate that the operators change sign when exchanged: $c_\alpha^\dagger c_\beta^\dagger = -c_\beta^\dagger c_\alpha^\dagger$. This is more conveniently written as $\{c_\alpha^\dagger, c_\beta^\dagger\} = 0$ by introducing the *anti-commutator*

$$\{A, B\} := A\,B + B\,A\,. \tag{67}$$

The simplest state we can produce with these operators is the single-electron state $|\alpha\rangle = c_\alpha^\dagger|0\rangle$. When we want to calculate its norm, we have to consider the adjoint of $c_\alpha^\dagger|0\rangle$, formally obtaining $\langle\alpha|\alpha\rangle = \langle 0|c_\alpha c_\alpha^\dagger|0\rangle$, or, more generally, $\langle\alpha|\beta\rangle = \langle 0|c_\alpha c_\beta^\dagger|0\rangle$. This must mean that c_α, the adjoint of a creation operator, must remove an electron from the state, otherwise the overlap of $c_\alpha c_\beta^\dagger|0\rangle$ with the vacuum state $\langle 0|$ would vanish. We therefore call the adjoint of the creation operator an *annihilation operator*. We certainly cannot take an electron out of the vacuum state, so $c_\alpha|0\rangle = 0$. Moreover, by taking the adjoint or the anti-commutator of the creation operators, we see that also the annihilation operators anti-commute: $\{c_\alpha, c_\beta\} = 0$. Moreover, to obtain the proper normalization of the single-electron states, we postulate the commutation relation $\{c_\alpha, c_\beta^\dagger\} = \langle\alpha|\beta\rangle$.
Thus, we have defined the vacuum state $|0\rangle$ and the set of operators c_α related to single-electron states $|\alpha\rangle$ with the properties

$$\boxed{\begin{array}{ll} c_\alpha|0\rangle = 0 & \left\{c_\alpha, c_\beta\right\} = 0 = \left\{c_\alpha^\dagger, c_\beta^\dagger\right\} \\ \langle 0|0\rangle = 1 & \left\{c_\alpha, c_\beta^\dagger\right\} = \langle\alpha|\beta\rangle \end{array}} \tag{68}$$

We note that the creators and annihilators are not ordinary operators in a Hilbert space, but transfer states from an N-electron to a $N \pm 1$-electron Hilbert space, i.e., they are operators defined on the *Fock space*. It is also remarkable that the mixed anti-commutator is the only place where the orbitals that distinguish different operators enter.
One type of operators is particularly useful for making contact with the real-space picture: The operators $\hat\Psi^\dagger(x)$, with $x = (r, \sigma)$, that create an electron of spin σ at position r, i.e., in state $|x\rangle = |r, \sigma\rangle$. Because of their importance they get a special name, *field operators*, and a special symbol $\hat\Psi^\dagger(x)$ instead of $c_x^\dagger$, but really they are just ordinary creation operators for the states corresponding to a delta function at r and a spin σ. The anti-commutator for the field-operators obviously follow from (68)

$$\left\{\hat\Psi(x), \hat\Psi(x')\right\} = 0 = \left\{\hat\Psi^\dagger(x), \hat\Psi^\dagger(x')\right\} \quad \text{and} \quad \left\{\hat\Psi(x), \hat\Psi^\dagger(x')\right\} = \delta(x - x')\,. \tag{69}$$

Given the single-electron wave functions in real space $\varphi_\alpha(x)$, we can express any creation operator in terms of the field operators

$$c_\alpha^\dagger = \int dx\, \varphi_\alpha(x)\hat\Psi^\dagger(x)\,. \tag{70}$$

Using (69), it is easy to see that these operator indeed fulfill all properties (68) required of the creation operators.

Conversely, if we have a complete set of single electron states $\{\varphi_{\alpha_n}(x)\}$, we can expand the field operators in terms of the corresponding creators and annihilators. Given the overlap matrix $S = (\langle\alpha_n|\alpha_m\rangle)$ we can use the Cholesky factorization $S^{-1} = T^\dagger T$ to orthonormalize the orbitals $\tilde{\varphi}_{\alpha_n}(x) = \sum T_{n,m}\,\varphi_{\alpha_m}(x)$. The completeness relation is then

$$\sum_{n,m} \overline{\varphi_{\alpha_n}(x)}\ \left(S^{-1}\right)_{n,m}\ \varphi_{\alpha_m}(x') = \sum_j \overline{\tilde{\varphi}_{\alpha_j}(x)}\,\tilde{\varphi}_{\alpha_j}(x') = \delta(x-x')\,. \tag{71}$$

Using this together with the commutation relations (68) we see that the operators

$$\hat{\Psi}(x) = \sum_n \tilde{\varphi}_{\alpha_n}(x)\,c_{\alpha_n}\ , \tag{72}$$

fulfill the commutation relations (69) of the field operators.

4.2 Representation of Slater determinants

We now show that we can write a Slater determinant in terms of the algebra (68) we have just defined. For this we consider an N-electron state $\prod c^\dagger_\alpha\,|0\rangle$ and prove that its real-space representation, obtained via the field operators is just the corresponding Slater determinant

$$\Phi_{\alpha_1\alpha_2\ldots\alpha_N}(x_1,x_2,\ldots,x_N) = \frac{1}{\sqrt{N!}}\left\langle 0\,\middle|\,\hat{\Psi}(x_1)\hat{\Psi}(x_2)\ldots\hat{\Psi}(x_N)\,c^\dagger_{\alpha_N}\ldots c^\dagger_{\alpha_2}c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle \tag{73}$$

Not surprisingly, the proof is by induction. As a warm-up we consider the case of a single-electron wave function ($N=1$). Using the special case of an anti-commutation relation

$$\{\hat{\Psi}(x), c^\dagger_\alpha\} = \int dx'\,\varphi_\alpha(x')\left\{\hat{\Psi}(x),\hat{\Psi}^\dagger(x')\right\} = \varphi_\alpha(x) \tag{74}$$

we see that

$$\left\langle 0\,\middle|\,\hat{\Psi}(x_1)\,c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle = \left\langle 0\,\middle|\,\varphi_{\alpha_1}(x_1) - c^\dagger_{\alpha_1}\hat{\Psi}(x_1)\,\middle|\,0\right\rangle = \varphi_{\alpha_1}(x_1) \tag{75}$$

For the two-electron state $N=2$, we anticommute $\hat{\Psi}(x_2)$ in two steps to the right

$$\begin{aligned}
\left\langle 0\,\middle|\,\hat{\Psi}(x_1)\hat{\Psi}(x_2)\,c^\dagger_{\alpha_2}c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle &= \left\langle 0\,\middle|\,\hat{\Psi}(x_1)\left(\varphi_{\alpha_2}(x_2) - c^\dagger_{\alpha_2}\hat{\Psi}(x_2)\right)c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle \\
&= \left\langle 0\,\middle|\,\hat{\Psi}(x_1)c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle\,\varphi_{\alpha_2}(x_2) - \left\langle 0\,\middle|\,\hat{\Psi}(x_1)c^\dagger_{\alpha_2}\hat{\Psi}(x_2)c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle \\
&= \varphi_{\alpha_1}(x_1)\varphi_{\alpha_2}(x_2) - \varphi_{\alpha_2}(x_1)\varphi_{\alpha_1}(x_2)\,.
\end{aligned} \tag{76}$$

We see how anti-commutating automatically produces appropriate sign for the antisymmetric wave function. Dividing by $\sqrt{2}$, we obtain the desired two-electron Slater determinant.
The general case of an N-electron state works just the same. Anti-commuting $\hat{\Psi}(x_N)$ all the way to the right produces $N-1$ terms with alternating sign

$$\begin{aligned}
&\left\langle 0\,\middle|\,\hat{\Psi}(x_1)\ldots\hat{\Psi}(x_{N-1})\hat{\Psi}(x_N)\,c^\dagger_{\alpha_N}c^\dagger_{\alpha_{N-1}}\ldots c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle\ = \\
&\quad+\quad \left\langle 0\,\middle|\,\hat{\Psi}(x_1)\ldots\hat{\Psi}(x_{N-1})\,c^\dagger_{\alpha_{N-1}}\ldots c^\dagger_{\alpha_1}\,\middle|\,0\right\rangle \quad \varphi_{\alpha_N}(x_N) \\
&\quad-\quad \left\langle 0\,\middle|\,\hat{\Psi}(x_1)\ldots\hat{\Psi}(x_{N-1})\,\textstyle\prod_{n\neq N-1}c^\dagger_{\alpha_n}\,\middle|\,0\right\rangle \quad \varphi_{\alpha_{N-1}}(x_N) \\
&\quad\vdots \\
&(-1)^N\quad \left\langle 0\,\middle|\,\hat{\Psi}(x_1)\ldots\hat{\Psi}(x_{N-1})\,c^\dagger_{\alpha_N}\ \ldots\ c^\dagger_{\alpha_2}\,\middle|\,0\right\rangle \quad \varphi_{\alpha_1}(x_N)
\end{aligned}$$

Using (73) for the $N-1$-electron states, this is nothing but the Laplace expansion of

$$D = \begin{vmatrix} \varphi_{\alpha_1}(x_1) & \varphi_{\alpha_2}(x_1) & \cdots & \varphi_{\alpha_N}(x_1) \\ \varphi_{\alpha_1}(x_2) & \varphi_{\alpha_2}(x_2) & \cdots & \varphi_{\alpha_N}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(x_N) & \varphi_{\alpha_2}(x_N) & \cdots & \varphi_{\alpha_N}(x_N) \end{vmatrix}$$

along the Nth row. Dividing by $\sqrt{N!}$ we see that we have shown (73) for N-electron states.
Thus we see that instead of working with the Slater determinant $\Phi_{\alpha_1\alpha_2\ldots\alpha_N}(x_1, x_2, \ldots, x_N)$ we can work with the corresponding N-electron product state $\prod c^\dagger_\alpha |0\rangle$. In particular, instead of working with the basis of Slater determinants

$$\left\{\Phi_{\alpha_1\cdots\alpha_N}(x_1,\ldots,x_N) \,\middle|\, \alpha_1 < \alpha_2 < \cdots < \alpha_N \in \{1,\ldots,K\}\right\} \tag{77}$$

induced by an orthonormal set of single-electron states $\{\varphi_{\alpha_n}(x)\}$, we can work with the corresponding basis of product states

$$\left\{ \prod_{\alpha_1<\cdots<\alpha_N} c^\dagger_{\alpha_N}\cdots c^\dagger_{\alpha_1}|0\rangle \right\}. \tag{78}$$

4.3 Representation of n-body operators

By the definition of indistinguishability, observables on an N-particle Hilbert space must not distinguish between the N indistinguishable particles. Hence they must be symmetric under particle permutation. We can then write a general operator as a sum of n-particle operators

$$M(\mathbf{x}) = M_0 + \sum_i M_1(x_i) + \sum_{i<j} M_2(x_i,x_j) + \sum_{i<j<k} M_3(x_i,x_j,x_k) + \cdots \tag{79}$$

where the summations can be restricted since the operators must be symmetric in their arguments, e.g., $M_2(x_1,x_2) = M(x_2,x_1)$.
Having established the relation between product states and Slater determiants, it is straightforward to express the matrix elements of a general n-body operator with N-particle Slater determinants:

$$\begin{aligned}\int dx_1\cdots dx_N\, \overline{\Phi_{\beta_1\cdots\beta_N}(x_1,\cdots,x_N)} M(x_1,\cdots,x_N)\Phi_{\alpha_1\cdots\alpha_N}(x_1,\cdots,x_N) \\ = \left\langle 0 \,\middle|\, c_{\beta_1}\cdots c_{\beta_N}\, \hat{M}\, c^\dagger_{\alpha_N}\cdots c^\dagger_{\alpha_1} \,\middle|\, 0\right\rangle\end{aligned} \tag{80}$$

with the representation of the n-body operator in terms of field operators

$$\hat{M} = \frac{1}{N!}\int dx_1\cdots x_N\, \hat\Psi^\dagger(x_N)\cdots\hat\Psi^\dagger(x_1)\, M(x_1,\cdots,x_N)\, \hat\Psi(x_1)\cdots\hat\Psi(x_N)\,. \tag{81}$$

Note that this form of the operator is only valid when applied to N-electron states. But from here on, we can work entirely in terms of our algebra (68).

To see what (81) means we look at its parts (79). As usual, we start with the simplest case, the zero-body operator, which, up to trivial prefactor, is $M_0(x_1, \cdots, x_N) = 1$. Operating on an N-electron wave function, it gives

$$\begin{aligned}
\hat{M}_0 &= \frac{1}{N!} \int dx_1 dx_2 \cdots x_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \hat{\Psi}^\dagger(x_1) \, \hat{\Psi}(x_1) \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\
&= \frac{1}{N!} \int \quad dx_2 \cdots x_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \qquad \hat{N} \qquad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\
&= \frac{1}{N!} \int \quad dx_2 \cdots x_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \qquad 1 \qquad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\
&\vdots \\
&= \frac{1}{N!} \, 1 \cdot 2 \, \cdots \, N = 1
\end{aligned} \tag{82}$$

where we have used that

$$\int dx \, \hat{\Psi}^\dagger(x) \hat{\Psi}(x) = \hat{N} \tag{83}$$

is the number operator and that applying n annihilation operators $\hat{\Psi}(x_j)$ to an N-electron state gives a state with $N - n$ electrons. We note that we obtain a form of $\hat{M}_0 = 1$ that apparently does not depend on the number of electrons in the wave function that it is applied to. This was not the case for the original expression (81).
Next we consider one-body operators $M(x_1, \ldots, x_N) = \sum_j M_1(x_j)$

$$\begin{aligned}
\hat{M}_1 &= \frac{1}{N!} \int dx_1 \cdots dx_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_1) \sum_j M_1(x_j) \, \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N) \\
&= \frac{1}{N!} \sum_j \int dx_j \, \hat{\Psi}^\dagger(x_j) \, M_1(x_j) \, (N-1)! \, \hat{\Psi}(x_j) \\
&= \frac{1}{N} \sum_j \int dx_j \, \hat{\Psi}^\dagger(x_j) \, M_1(x_j) \, \hat{\Psi}(x_j) \\
&= \qquad \int dx \quad \hat{\Psi}^\dagger(x) \;\; M_1(x) \;\; \hat{\Psi}(x)
\end{aligned}$$

Here we have first anticommuted $\hat{\Psi}^\dagger(x_j)$ all the way to the left and $\hat{\Psi}(x_j)$ to the right. Since these take the same numbers of anticommutations, there is no sign involved. The operation leaves the integrals over the variables except x_i, a zero-body operator for $N - 1$ electron states, operating on $\hat{\Psi}(x_j)|N\text{-electron state}\rangle$.
Expanding the field-operators in a complete orthonormal set $\hat{\Psi}(x) = \sum_n \varphi_{\alpha_n}(x) \, c_{\alpha_n}$ gives

$$\hat{M}_1 = \sum_{n,m} \int dx \, \overline{\varphi_{\alpha_n}(x)} \, M(x) \, \varphi_{\alpha_m}(x) \; c^\dagger_{\alpha_n} c_{\alpha_m} = \sum_{n,m} \langle \alpha_n | M_1 | \alpha_m \rangle \; c^\dagger_{\alpha_n} c_{\alpha_m} \; . \tag{84}$$

Also here we find a form for $\hat{M}_1$ that is apparently independent of the number of electrons N and can be evaluated directly in the basis states (78).
For the two-body operators $M(x_1, \ldots, x_N) = \sum_{i<j} M_2(x_i, x_j)$ we proceed in the familiar way, anti-commuting first the operators with the coordinates involved in M_2 all the way to the left

and right. This time we are left with a zero-body operator for $N-2$ electrons:

$$\begin{aligned}
\hat{M}_2 &= \frac{1}{N!}\int dx_1\cdots dx_N\,\hat\Psi^\dagger(x_N)\cdots\hat\Psi^\dagger(x_1)\sum_{i<j}M_2(x_i,x_j)\;\hat\Psi(x_1)\cdots\hat\Psi(x_N)\\
&= \frac{1}{N!}\sum_{i<j}\int dx_i dx_j\,\hat\Psi^\dagger(x_j)\hat\Psi^\dagger(x_i)\,M_2(x_i,x_j)\,(N-2)!\;\hat\Psi(x_i)\hat\Psi(x_j)\\
&= \frac{1}{N(N-1)}\sum_{i<j}\int dx_i dx_j\,\hat\Psi^\dagger(x_j)\hat\Psi^\dagger(x_i)\,M_2(x_i,x_j)\,\hat\Psi(x_i)\hat\Psi(x_j)\\
&= \frac{1}{2}\int dx\,dx'\;\hat\Psi^\dagger(x')\;\hat\Psi^\dagger(x)\;\;M_2(x,x')\;\;\hat\Psi(x)\;\hat\Psi(x')
\end{aligned}$$

Expanding in an orthonormal basis, we get

$$\begin{aligned}
\hat{M}_2 &= \frac{1}{2}\sum_{n,n',m,m'}\int dx dx'\,\overline{\varphi_{\alpha_{n'}}(x')\varphi_{\alpha_n}(x)}\,M_2(x,x')\,\varphi_{\alpha_m}(x)\varphi_{\alpha_{m'}}(x')\quad c^\dagger_{\alpha_{n'}}c^\dagger_{\alpha_n}c_{\alpha_m}c_{\alpha_{m'}}\\
&= \frac{1}{2}\sum_{n,n',m,m'}\qquad\qquad\langle\alpha_n\alpha_{n'}|M_2|\alpha_m\alpha_{m'}\rangle\qquad\qquad c^\dagger_{\alpha_{n'}}c^\dagger_{\alpha_n}c_{\alpha_m}c_{\alpha_{m'}}
\end{aligned}\tag{85}$$

where the exchange of the indices in the second line is a consequence of the way the Dirac state for two electrons is usually written: first index for the first coordinate, second index for the second, while taking the adjoint of the operators changes their order. Obviously, from the symmetry $M_2(x,x') = M_2(x',x)$ follows $\langle\alpha_n\alpha_{n'}|M_2|\alpha_m\alpha_{m'}\rangle = \langle\alpha_{n'}\alpha_n|M_2|\alpha_{m'}\alpha_m\rangle$.
The procedure generalizes to operators acting on more than two electrons in the natural way.
We note that, while we started from a form of the operators (79) that was explicitly formulated in an N-electron Hilbert space, the results (82), (84), and (85) are of the same form no matter what value N takes. Thus these operators are valid not just on some N-electron Hilbert space, but on the entire Fock space. This is a particular strength of the second quantized formulation.
For a more in-depth discussion of second quantization and the consequences of the indistinguishability of elementary particles, see [22], form which the present section was taken.

Acknowledgment

Support of the Deutsche Forschungsgemeinschaft through FOR1346 is gratefully acknowledged.

Appendices

A Atomic units

Practical electronic structure calculations are usually done in atomic units, a. u. for short. While the idea behind the atomic units is remarkably simple, in practice there is often some confusion when trying to convert to SI units. We therefore give a brief explanation.
The motivation for introducing atomic units is to simplify the equations. For example, in SI units the Hamiltonian of a hydrogen atom is

$$H = -\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2}{4\pi\epsilon_0\, r}\ . \tag{86}$$

To avoid having to keep track of the constants, we would like to simplify this to

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r}\ . \tag{87}$$

To this end we invent units in which the *numerical values* of the electron mass m_e, the elementary charge e, the Planck-constant $\hbar$, and the dielectric constant $4\pi\epsilon_0$ are all equal to one. This immediately tells us: 1 a.u. mass $= m_e$ and 1 a.u. charge $= e$. To complete the set of basis units we still need the atomic unit of length, which we call a_0, and of time, t_0. To find the values of a_0 and t_0 we write $\hbar$ and $4\pi\epsilon_0$ (using simple dimensional analysis) in atomic units: $\hbar = 1\, m_e a_0^2/t_0$ and $4\pi\epsilon_0 = 1\, t_0^2 e^2/(m_e a_0^3)$. Solving this system of equations, we find

1 a.u. length	$= a_0 =$	$4\pi\epsilon_0\hbar^2/m_e e^2$	$\approx 5.2918\cdot 10^{-11}$	m
1 a.u. mass	$= m_e =$		$\approx 9.1095\cdot 10^{-31}$	kg
1 a.u. time	$= t_0 =$	$(4\pi\epsilon_0)^2\hbar^3/m_e e^4$	$\approx 2.4189\cdot 10^{-17}$	s
1 a.u. charge	$= e =$		$\approx 1.6022\cdot 10^{-19}$	C

The atomic unit of length, a_0, is the Bohr radius. As the dimension of energy is mass times length squared divided by time squared, its atomic unit is $m_e a_0^2/t_0^2 = m_e e^4/(4\pi\epsilon_0)^2\hbar^2$. Because of its importance the atomic unit of energy has a name, the Hartree. One Hartree is minus twice the ground-state energy of the hydrogen atom (86), about 27.211 eV. The speed of light c in atomic units is given by $c\ t_0/a_0 = 4\pi\varepsilon_0\hbar c/e^2 = 1/\alpha$, where α is the fine structure constant. Thus $c = \alpha^{-1}$ a.u. ≈ 137 a.u. The Bohr magneton is $\mu_B = 1/2$ a.u.

B Downfolding

To integrate-out high-energy degrees of freedom, we partition the Hilbert space of the full system into states of interest (low-energy states) and 'other' states, which will be integrated out. The Hamiltonian is then written in blocks

$$H = \begin{pmatrix} H_{00} & T_{01} \\ T_{10} & H_{11} \end{pmatrix} , \tag{88}$$

where H_{00} is the Hamiltonian restricted to the states of interest (reduced Hilbert space), H_{11} the Hamiltonian for the 'other' states, and the T matrices describe transitions between the two subspaces. The resolvent is partitioned likewise

$$G(\varepsilon) = (\varepsilon - H)^{-1} = \begin{pmatrix} \varepsilon - H_{00} & T_{01} \\ T_{10} & \varepsilon - H_{11} \end{pmatrix}^{-1} . \tag{89}$$

Calculating the inverse of the 2×2 matrix, taking into account that the entries are matrices themselves and thus do not commute, we obtain

$$G_{00}(\varepsilon) = \Bigg(\varepsilon - [H_{00} + \underbrace{T_{01}(\varepsilon - H_{11})^{-1} T_{10}}_{=H_{\text{eff}}}] \Bigg)^{-1} . \tag{90}$$

This expression looks just like the resolvent for a Hamiltonian

$$\begin{aligned} H_{\text{eff}} &= H_{00} + T_{01}(\varepsilon - H_{11})^{-1} T_{10} && (91) \\ &\approx H_{00} + T_{01}(\varepsilon_0 - H_{11})^{-1} T_{10} && (92) \end{aligned}$$

on the reduced Hilbert space. This effective Hamiltonian describes the physics of the full system, but operates only on the small reduced Hilbert space. Of course, this drastic simplification comes at a price: the effective Hamiltonian is energy dependent. If the hopping matrix elements in T_{01} are small, and/or if the states in the part of the Hilbert space that has been integrated out are energetically well-separated from the states that are explicitly considered, this energy dependence can, to a good approximation, be neglected. We can then replace ε by a typical energy ε_0 for the states in the reduced Hilbert space to obtain an energy-independent Hamiltonian $H_{\text{eff}}(\varepsilon_0)$ that gives a good description of the electrons in the reduced Hilbert space, i.e., the states with an energy close to ε_0.

C Pauli matrices

Here we collect the most important properties of the Pauli matrices. The Pauli or spin matrices are defined as

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{93}$$

They are hermitean, i.e. $\sigma_i^\dagger = \sigma_i$, and $\sigma_i^2 = 1$. Therefore their eigenvalues are ± 1. The eigenvectors of σ_z are $|m_z\rangle$, $m_z = \pm 1$:

$$|+1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |-1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{94}$$

For these vectors we find

$$\sigma_x|\,m_z\rangle = |\,-m_z\rangle \quad \sigma_y|\,m_z\rangle = im_z|\,-m_z\rangle \quad \sigma_z|\,m_z\rangle = m_z|\,m_z\rangle \tag{95}$$

The products of the Pauli matrices are $\sigma_x\,\sigma_y = i\sigma_z$, where the indices can be permuted cyclically. From this follows for the commutator

$$[\sigma_x, \sigma_y] = 2i\sigma_z \tag{96}$$

while the anticommutator vanishes:

$$\{\sigma_x, \sigma_y\} = 0 \tag{97}$$

Finally a rotation by an angle ϕ about the axis $\hat{n}$ changes the spin matrices

$$R_{\hat{n}}(\phi) = e^{-i\hat{n}\cdot\vec{\sigma}\,\phi/2} = cos(\phi/2) - i\sin(\phi/2)\,\hat{n}\cdot\vec{\sigma} \tag{98}$$

References

[1] R. Peierls: *Surprises in Theoretical Physics* (Princeton University Press, 1979)

[2] J.H. van Vleck: *Quantum Mechanics: The Key to Understanding Magnetism*, in S. Lundqvist (ed.): *Nobel Lectures, Physics 1971-1980* (World Scientific, Singapore, 1992) http://www.nobelprize.org/nobel_prizes/physics/laureates/1977

[3] E. Pavarini, E. Koch, F. Anders, and M. Jarrell (eds.): *Correlated Electrons: From Models to Materials* Reihe Modeling and Simulation, Vol. 2 (Forschungszentrum Jülich, 2012) http://www.cond-mat.de/events/correl12

[4] E. Pavarini, E. Koch, D. Vollhardt, and A.I. Lichtenstein: *The LDA+DMFT approach to strongly correlated materials* Reihe Modeling and Simulation, Vol. 1 (Forschungszentrum Jülich, 2011) http://www.cond-mat.de/events/correl11

[5] E. Pavarini, E. Koch, and U. Schollwöck (eds): *Emergent Phenomena in Correlated Matter* Reihe Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013) http://www.cond-mat.de/events/correl13

[6] E. Pavarini: *The LDA+DMFT Approach*, in [4]

[7] Patrik Fazekas: *Lecture Notes on Electron Correlation and Magnetism* (World Scientific, Singapore, 1999)

[8] A. Auerbach: *Interacting Electrons and Quantum Magnetism* (Springer, Heidelberg, 1994)

[9] R. Eder: *Multiplets in Transition Metal Ions*, in [3]

[10] M. Weissbluth: *Atoms and Molecules* (Academic Press, San Diego, 1978)

[11] P.W. Anderson: *Theory of Magnetic Exchange Interactions*, in F. Seitz and D. Turnbull (eds.): *Solid State Physics* **14**, 99 (1963)

[12] D.C. Mattis: *The Theory of Magnetism Made Simple* (World Scientific, Singapore, 2006)

[13] S. Blundell: *Magnetism in Condensed Matter* (Oxford University Press, 2001)

[14] A. Szabo and N.S. Ostlund: *Modern Quantum Chemistry* (Dover Publications, 1996)

[15] J.B. Goodenough: *Goodenough-Kanamori rule*, Scholarpedia **3**, 7382 (2008) http://scholarpedia.org/article/Goodenough-Kanamori_rule

[16] A. Montorsi (Ed.): *The Hubbard Model*, (World Scientific, Singapore, 1992)

[17] F. Gebhard: *The Mott Metal-Insulator Transition* (Springer, Heidelberg, 1997)

[18] A. Messiah: *Méchanique Quantique* (Dunond, Paris, 1964)

[19] P.M.A. Dirac: *The Quantum Theory of Emission and Absorbtion of Radiation*, Proc. Roy. Soc. A **114**, 243 (1927)

[20] P. Jordan and O. Klein: *Zum Mehrkörperproblem in der Quantenmechanik*, Z. Physik **45**, 751 (1927)

[21] P. Jordan and E. Wigner: *Über das Paulische Äquivalenzverbot*, Z. Physik **47**, 631 (1928)

[22] E. Koch: *Many-Electron States*, in [5]

B 2 Keldysh formalism for nonequilibrium transport through quantum systems[1]

S. G. Jakobs
Institut für Theorie der Statistischen Physik
RWTH Aachen University

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

The Keldysh formalism allows to analyze quantum mechanical nonequilibrium situations. In this lecture we will introduce basic concepts of this formalism and use them to describe transport through quantum systems. The typical setup we have in mind consists of a quantum dot or wire which is coupled to several leads. Interesting observables are for example the electric current and the heat current through the quantum dot.

Over the last decades the progress in micro-fabrication technologies has stimulated much research on transport properties of such mesoscopic systems. Apart from prospects of future applications in information technology, the investigation of quantum dots and wires provides an excellent possibility to enhance the understanding of particle correlations. Due to their smallness mesoscopic systems are characterized by full or partial coherence of the quantum mechanical phase. Interactions can have drastic influence on their properties. Examples for interaction induced phenomena are the Luttinger liquid behaviour of one-dimensional interacting quantum systems or the Kondo effect induced by a magnetic impurity in a conductor. In contrast to studies on condensed matter bulk materials, mesoscopic quantum systems allow to explore the properties of a single isolated structure, to tailor its geometry, to tune and measure many important parameters, and even to transfer it into a non-thermal state by applying a local bias voltage or temperature gradient.

Keldysh formalism describes the time evolution of observables after a system has been prepared in an initial density matrix. The calculation of expectation values involves a forward and a backward time evolution corresponding to the evolution of "the ket and the bra part" of the density matrix. The existence of these two time evolutions is responsible for a tensor structure of the Green and vertex functions. This makes Keldysh formalism more complex than, for instance, the Matsubara formalism which is widely used to study equilibrium situations. However, the real-time structure of Keldysh formalism is connected to two advantages. First, the formalism allows to describe non-equilibrium situations like transport at finite bias voltage. Second, dynamic quantities are readily found in their dependence on real times or frequencies. In Matsubara formalism in contrast an analytic continuation from the imaginary to the real axis is required, which can be numerically quite nontrivial.

This lecture consists of two parts. The first part (section 2) introduces general features of Keldysh formalism without special focus on mesoscopic transport setups. We consider the time evolution of a rather generic interacting quantum system and establish the concept of the time loop contour. Expectation values and correlation functions can be derived from suitably defined Green functions. We derive a diagrammatic language for expanding those in powers of the interaction. As important linear transformations we study the Keldysh rotation and, for stationary state situations, the Fourier transformation from time to frequency arguments. We mention which properties of the Green functions are characteristic for thermal equilibrium and conclude the part by indicating briefly how properties of Green and vertex functions can be understood from the generating functional perspective. The scope of this lecture allows to address only a selection of basic properties of the formalism. More information on Keldysh formalism can be found for instance in [1, 2, 3, 4, 5].

The second part (section 3) deals with nonequilibrium transport through quantum systems. We introduce a (still rather generic) model for a quantum dot coupled to leads and describe how the influence of the leads on the dot can be described by a self-energy contribution to the dot propagation. We derive how dot occupancy, particle current and heat current can be computed from Green functions and discuss transport coefficients characterizing the features of the sys-

tem in linear response. We do not evaluate finally our equations for specific models. This would require explicit results for the interacting Green functions. It is a major task to find those in reasonable approximations. A first approach is given by perturbation theory in the interaction (for which we derive the diagram rules). Its applicability has however limitations. In low dimensional Fermi systems for instance, perturbation theory can lead to divergencies. A more sophisticated approach is provided for example by renormalization group methods like the functional renormalization group, see lecture B7 by V. Meden. An analysis of specific quantum impurity models with the numerical renormalization group is presented in lecture B3 by T. A. Costi.
Throughout the present lecture we use units in which $\hbar = 1$ and $k_{\mathrm{B}} = 1$.

2 Keldysh formalism and diagrammatics

2.1 System under consideration

Consider a quantum system of which we know the statistical operator at some instant in time,

$$\hat{\rho}(t_0) = \hat{\rho}_0, \tag{1}$$

and the Hamiltonian $\hat{H}(t)$. For $t > t_0$ the Hamiltonian induces a unitary time evolution of the system. Keldysh formalism is designed to study expectation values of operators or operator-operator correlation functions for times $t > t_0$.
In order to introduce the formalism in this section, we consider a system of a single kind of identical particles with the Hamiltonian

$$\hat{H}(t) = \hat{H}_0(t) + \hat{V}(t) \tag{2}$$

with

$$\hat{H}_0(t) = \sum_{s',s} \epsilon_{s's}(t) \hat{a}^\dagger_{s'} \hat{a}_s, \tag{3}$$

$$\hat{V}(t) = \frac{1}{4} \sum_{\substack{s'_1,s_1 \\ s'_2,s_2}} \overline{v}_{s'_1 s'_2 s_1 s_2}(t) \hat{a}^\dagger_{s'_1} \hat{a}^\dagger_{s'_2} \hat{a}_{s_2} \hat{a}_{s_1}, \tag{4}$$

in standard notation of second quantization. Here, s denotes states of an orthonormal basis of the single-particle Hilbert space. The annihilator $\hat{a}_s$ is either Fermionic or Bosonic,

$$[\hat{a}_s, \hat{a}^\dagger_{s'}]_{-\zeta} = \delta_{ss'}, \tag{5}$$

$$\zeta = \begin{cases} 1, & \text{for Bosons,} \\ -1, & \text{for Fermions.} \end{cases} \tag{6}$$

The single-particle matrix elements of the Hamiltonian are

$$\epsilon_{s's}(t) = \langle s' | \hat{H}_0(t) | s \rangle \tag{7}$$

while

$$\overline{v}_{s'_1 s'_2 s_1 s_2}(t) = \langle s'_1 s'_2 | \hat{V}(t) | s_1 s_2 \rangle + \zeta \langle s'_1 s'_2 | \hat{V}(t) | s_2 s_1 \rangle \tag{8}$$

denotes the (anti-)symmetrized matrix elements of the two-particle interaction.
The formalism can as well treat mixtures of different particle types with different kinds of interactions. We restrict this exposition to the rather generic model with two-particle interaction for the sake of transparency. In section 3.1, the model will be specified further in order to discuss transport through mesoscopic systems.

2.2 Contour ordered Green functions

Time loop contour. Let us first analyze the expectation value of some operator $\hat{A}(t)$,

$$\langle \hat{A} \rangle(t) = \mathrm{Tr}\, \hat{A}(t)\hat{\rho}(t) = \mathrm{Tr}\, \hat{A}(t)_{\mathrm{H}}\hat{\rho}_0 = \mathrm{Tr}\, \hat{u}(t_0,t)\hat{A}(t)\hat{u}(t,t_0)\hat{\rho}_0, \tag{9}$$

where

$$\hat{A}(t)_{\mathrm{H}} = \hat{u}(t_0,t)\hat{A}(t)\hat{u}(t,t_0) \tag{10}$$

is the operator in the Heisenberg picture with reference time t_0. The time evolution operator $\hat{u}(t,t')$ is determined by

$$i\partial_t \hat{u}(t,t') = \hat{H}(t)\hat{u}(t,t') \tag{11}$$

and the group property

$$\hat{u}(t,t')\hat{u}(t',t'') = \hat{u}(t,t''). \tag{12}$$

It is given by

$$\hat{u}(t,t') = \begin{cases} \sum_{n=0}^{\infty}(-i)^n \int_{t>\tau_n>\cdots>\tau_1>t'} d\tau_n \ldots d\tau_1 \hat{H}(\tau_n)\ldots\hat{H}(\tau_1), & t > t', \\ \sum_{n=0}^{\infty} i^n \int_{t<\tau_n<\cdots<\tau_1<t'} d\tau_n \ldots d\tau_1 \hat{H}(\tau_n)\ldots\hat{H}(\tau_1), & t < t' \end{cases} \tag{13}$$

$$= \begin{cases} \hat{\mathcal{T}} \exp\left[-i\int_{t'}^{t} \hat{H}(\tau)d\tau\right], & t > t', \\ \hat{\tilde{\mathcal{T}}} \exp\left[i\int_{t}^{t'} \hat{H}(\tau)d\tau\right], & t < t'. \end{cases} \tag{14}$$

Here, $\hat{\mathcal{T}}$ is the time ordering operator which rearranges a product of operators such that time arguments increase from right to left. The anti-time ordering operator $\hat{\tilde{\mathcal{T}}}$ establishes the opposite order. For the expectation value follows

$$\langle \hat{A} \rangle(t) = \mathrm{Tr}\left[\hat{\tilde{\mathcal{T}}} e^{i\int_{t_0}^{t}\hat{H}(\tau)d\tau}\right] \hat{A}(t) \left[\hat{\mathcal{T}} e^{-i\int_{t_0}^{t}\hat{H}(\tau)d\tau}\right] \hat{\rho}_0. \tag{15}$$

The forward and backward time evolution in this expression can be interpreted as two branches of a time loop contour which leads from t_0 to t and back. Let us introduce a contour index j for operators. It can take values $j = \mathrm{f}$ (forward) and $j = \mathrm{b}$ (backward) and indicates, which branch the operator is associated with. Furthermore, we define the action of the contour ordering operator $\hat{\mathcal{T}}_{\mathrm{c}}$ on a product of operators as follows: all operators with index $j = \mathrm{b}$ are placed left of all operators with index $j = \mathrm{f}$; the block of b-operators is anti-time ordered internally, while the block of f-operators is time ordered. In total, the operators are sorted in increasing time loop order from right to left, compare Fig. 1. The expectation value can now be cast into the form

$$\langle \hat{A} \rangle(t) = \mathrm{Tr}\, \hat{\mathcal{T}}_{\mathrm{c}}\, e^{i\int_{t_0}^{t}\hat{H}^{\mathrm{b}}(\tau)d\tau} \hat{A}^{\mathrm{f}}(t) e^{-i\int_{t_0}^{t}\hat{H}^{\mathrm{f}}(\tau)d\tau}\hat{\rho}_0 \tag{16}$$

$$= \mathrm{Tr}\, \hat{\mathcal{T}}_{\mathrm{c}}\, e^{-i\int_{t_0}^{\infty}[\hat{H}^{\mathrm{f}}(\tau)-\hat{H}^{\mathrm{b}}(\tau)]d\tau} \hat{A}^{\mathrm{f}}(t)\hat{\rho}_0. \tag{17}$$

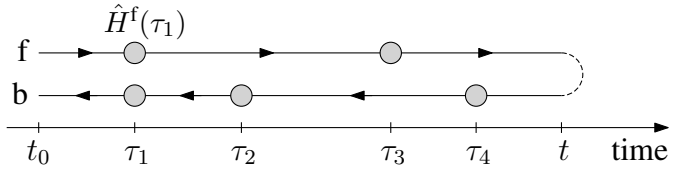


Fig. 1: *The time loop contour consists of a forward and a backward branch and leads from t_0 to t and back. An example for the action of the contour ordering operator is $\hat{\mathcal{T}}_{\mathrm{c}}\,\hat{H}^{\mathrm{f}}(\tau_1)\hat{H}^{\mathrm{b}}(\tau_1)\hat{H}^{\mathrm{b}}(\tau_2)\hat{H}^{\mathrm{f}}(\tau_3)\hat{H}^{\mathrm{b}}(\tau_4) = \hat{H}^{\mathrm{b}}(\tau_1)\hat{H}^{\mathrm{b}}(\tau_2)\hat{H}^{\mathrm{b}}(\tau_4)\hat{H}^{\mathrm{f}}(\tau_3)\hat{H}^{\mathrm{f}}(\tau_1)$.*

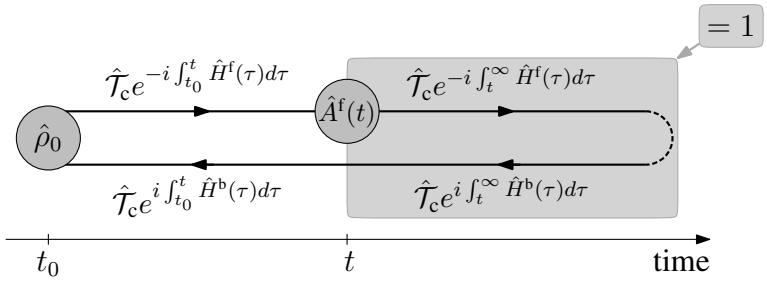


Fig. 2: *Graphical representation of $\mathrm{Tr}\,\hat{\mathcal{T}}_{\mathrm{c}}\, e^{-i\int_{t_0}^{\infty}[\hat{H}^{\mathrm{f}}(\tau)-\hat{H}^{\mathrm{b}}(\tau)]d\tau}\hat{A}^{\mathrm{f}}(t)\hat{\rho}_0$. The time evolution from t to ∞ and back cancels out. The operator $\hat{A}(t)$ could as well be placed on the backward contour.*

For the last step we introduced

$$1 = \hat{u}(t,\infty)\hat{u}(\infty,t) = \hat{\mathcal{T}}_{\mathrm{c}}\, e^{i\int_{t}^{\infty}\hat{H}^{\mathrm{b}}(\tau)d\tau} e^{-i\int_{t}^{\infty}\hat{H}^{\mathrm{f}}(\tau)d\tau}. \tag{18}$$

We assigned the index $j = \mathrm{f}$ to $\hat{A}(t)$ but could as well have chosen the index $j = \mathrm{b}$. Fig. 2 shows a graphical representation of the result for $\langle\hat{A}\rangle(t)$.
For the correlation function of two operators we get in complete analogy

$$\langle\hat{A}\hat{B}\rangle(t,t') = \mathrm{Tr}\,\hat{A}(t)_{\mathrm{H}}\hat{B}(t')_{\mathrm{H}}\hat{\rho}_0 = \mathrm{Tr}\,\hat{\mathcal{T}}_{\mathrm{c}}\, e^{-i\int_{t_0}^{\infty}[\hat{H}^{\mathrm{f}}(\tau)-\hat{H}^{\mathrm{b}}(\tau)]d\tau}\hat{A}^{\mathrm{b}}(t)\hat{B}^{\mathrm{f}}(t')\rho_0. \tag{19}$$

Here, we assigned $\hat{A}$ to the backward contour and $\hat{B}$ to the forward contour, such that contour order keeps $\hat{A}$ left of $\hat{B}$, irrespective of the time arguments t, t'. In order to express correlation functions of three or more operators in a similar way, additional branches of the contour would be required.

Green functions. So far, $\hat{\mathcal{T}}_{\mathrm{c}}$ merely shifted the parts of the time evolution to their appropriate positions. In order to define contour ordered Green functions, we need to generalize the action of $\hat{\mathcal{T}}_{\mathrm{c}}$ to arbitrary products of generators and annihilators with individual time and contour arguments. We agree again that $\hat{\mathcal{T}}_{\mathrm{c}}$ establishes increasing time loop order from right to left. If some operators from the same branch of the contour have identical time arguments, then the creators are sorted to the left of the annihilators. The result is multiplied by ζ if it is an odd permutation of the initial order. The contour ordered n-particle Green function is now defined as

$$\begin{aligned}
&G(x_1,\ldots,x_n|x_1',\ldots,x_n') \\
&= G^{j_1,\ldots,j_n|j_1',\ldots,j_n'}_{s_1,\ldots,s_n|s_1',\ldots,s_n'}(t_1,\ldots,t_n|t_1',\ldots,t_n') && (20)\\
&= (-i)^n\,\mathrm{Tr}\,\hat{\mathcal{T}}_{\mathrm{c}}\,\hat{a}^{j_1}_{s_1}(t_1)_{\mathrm{H}}\ldots\hat{a}^{j_n}_{s_n}(t_n)_{\mathrm{H}}\hat{a}^{\dagger\,j_n'}_{s_n'}(t_n')_{\mathrm{H}}\ldots\hat{a}^{\dagger\,j_1'}_{s_1'}(t_1')_{\mathrm{H}}\hat{\rho}_0 && (21)\\
&= (-i)^n\,\mathrm{Tr}\,\hat{\mathcal{T}}_{\mathrm{c}}\, e^{-i\int_{t_0}^{\infty}[\hat{H}^{\mathrm{f}}(\tau)-\hat{H}^{\mathrm{b}}(\tau)]d\tau}\hat{a}^{j_1}_{s_1}(t_1)\ldots\hat{a}^{j_n}_{s_n}(t_n)\hat{a}^{\dagger\,j_n'}_{s_n'}(t_n')\ldots\hat{a}^{\dagger\,j_1'}_{s_1'}(t_1')\hat{\rho}_0. && (22)
\end{aligned}$$

Here, $x_i = (j_i, t_i, s_i)$ is a multi-index. In the last expression, the time arguments of the creation and annihilation operators are just formal ones. They indicate the positions of the operators under contour ordering. Expectation values and correlation functions like $\operatorname{Tr} \hat{A}(t)_{\mathrm{H}} \hat{\rho}_0$ and $\operatorname{Tr} \hat{A}(t)_{\mathrm{H}} \hat{B}(t')_{\mathrm{H}} \hat{\rho}_0$ are superpositions of appropriate many-particle Green functions. This follows when $\hat{A}$ and $\hat{B}$ are expressed in second quantization. For instance, in section 3.3 we describe how the current through a quantum dot can be computed from appropriate Green functions.

The single-particle Green function has two contour indices. The corresponding $2 \times 2 = 4$ components are called lesser, greater, chronological, and anti-chronological Green function. Often it is handy to organize them in a matrix,

$$G^{j|j'} = \begin{pmatrix} G^{\mathrm{f|f}} & G^{\mathrm{f|b}} \\ G^{\mathrm{b|f}} & G^{\mathrm{b|b}} \end{pmatrix}_{jj'} = \begin{pmatrix} G^{\mathrm{c}} & G^{<} \\ G^{>} & G^{\tilde{\mathrm{c}}} \end{pmatrix}_{jj'}. \tag{23}$$

The components are given by

$$G^{\mathrm{f|b}}_{s|s'}(t|t') = G^{<}_{s|s'}(t|t') = -\zeta i \operatorname{Tr} \hat{a}^{\dagger}_{s'}(t')_{\mathrm{H}} \hat{a}_s(t)_{\mathrm{H}} \rho_0, \tag{24}$$

$$G^{\mathrm{b|f}}_{s|s'}(t|t') = G^{>}_{s|s'}(t|t') = -i \operatorname{Tr} \hat{a}_s(t)_{\mathrm{H}} \hat{a}^{\dagger}_{s'}(t')_{\mathrm{H}} \rho_0, \tag{25}$$

$$G^{\mathrm{f|f}}(t|t') = G^{\mathrm{c}}(t|t') = \begin{cases} G^{<}(t|t'), & t \leq t', \\ G^{>}(t|t'), & t > t', \end{cases} \tag{26}$$

$$G^{\mathrm{b|b}}(t|t') = G^{\tilde{\mathrm{c}}}(t|t') = \begin{cases} G^{>}(t|t'), & t < t', \\ G^{<}(t|t'), & t \geq t'. \end{cases} \tag{27}$$

They satisfy

$$G^{\mathrm{c}}(t|t') + G^{\tilde{\mathrm{c}}}(t|t') - G^{<}(t|t') - G^{>}(t|t') = \begin{cases} 0, & t \neq t', \\ i, & t = t'. \end{cases} \tag{28}$$

This trivial result highlights a redundancy among the four single-particle Green functions, which is exploited in the so called Keldysh rotation discussed in section 2.6.

Free Green functions. The free contour ordered Green functions are important for the diagrammatic expansion described below. For their definition we need the free time evolution operator $\hat{u}_0(t, t')$ as solution of

$$i\partial_t \hat{u}_0(t, t') = \hat{H}_0(t) \hat{u}_0(t, t'). \tag{29}$$

An operator in the interaction picture is given by

$$\hat{A}(t)_0 = \hat{u}_0(t_0, t) \hat{A}(t) \hat{u}_0(t, t_0), \tag{30}$$

and the free contour ordered n-particle Green function is

$$\begin{aligned} &g(x_1, \ldots, x_n | x'_1, \ldots, x'_n) \\ &= (-i)^n \operatorname{Tr} \hat{\mathcal{T}}_{\mathrm{c}} \, \hat{a}^{j_1}_{s_1}(t_1)_0 \ldots \hat{a}^{j_n}_{s_n}(t_n)_0 \hat{a}^{\dagger\, j'_n}_{s'_n}(t'_n)_0 \ldots \hat{a}^{\dagger\, j'_1}_{s'_1}(t'_1)_0 \hat{\rho}_0 \qquad (31) \\ &= (-i)^n \operatorname{Tr} \hat{\mathcal{T}}_{\mathrm{c}} \, e^{-i \int_{t_0}^{\infty} [\hat{H}^{\mathrm{f}}_0(\tau) - \hat{H}^{\mathrm{b}}_0(\tau)] d\tau} \hat{a}^{j_1}_{s_1}(t_1) \ldots \hat{a}^{j_n}_{s_n}(t_n) \hat{a}^{\dagger\, j'_n}_{s'_n}(t'_n) \ldots \hat{a}^{\dagger\, j'_1}_{s'_1}(t'_1) \hat{\rho}_0. \qquad (32) \end{aligned}$$

We compute the free single-particle Green functions explicitly. From

$$i\partial_t \hat{a}_s(t)_0 = i\partial_t \hat{u}_0(t_0,t)\hat{a}_s\hat{u}_0(t,t_0) = \hat{u}_0(t_0,t)[\hat{a}_s,\hat{H}_0(t)]_-\hat{u}_0(t,t_0), \tag{33}$$

and

$$[\hat{a}_s,\hat{H}_0(t)]_- = \sum_{r'r}\epsilon_{r'r}(t)[\hat{a}_s,\hat{a}^\dagger_{r'}\hat{a}_r]_- = \sum_r \epsilon_{sr}(t)\hat{a}_r \tag{34}$$

follows

$$i\partial_t g_{s|s'}(t|t') = \sum_r \epsilon_{sr}(t) g_{r|s'}(t|t'). \tag{35}$$

Similar steps show

$$i\partial_{t'} g_{s|s'}(t|t') = -\sum_{r'} g_{s|r'}(t|t')\epsilon_{r's'}(t'). \tag{36}$$

The initial values are given by

$$g^<_{s|s'}(t_0|t_0) = -i\zeta n_{ss'} \tag{37}$$

$$g^>_{s|s'}(t_0|t_0) = -i(\delta_{ss'} + \zeta n_{ss'}) \tag{38}$$

with the matrix of initial occupation

$$n_{ss'} = \mathrm{Tr}\,\hat{a}^\dagger_{s'}\hat{a}_s\hat{\rho}_0. \tag{39}$$

The solution is hence

$$g^<(t|t') = -i\zeta u_0(t,t_0)nu_0(t_0,t), \tag{40}$$

$$g^>(t|t') = -iu_0(t,t_0)(1+\zeta n)u_0(t_0,t), \tag{41}$$

$$g^{\mathrm{c}}(t|t') = \begin{cases} g^<(t|t'), & t \le t', \\ g^>(t|t'), & t > t', \end{cases} \tag{42}$$

$$g^{\tilde{\mathrm{c}}}(t|t') = \begin{cases} g^>(t|t'), & t < t', \\ g^<(t|t'), & t \ge t'. \end{cases} \tag{43}$$

Here, $u_0(t,t')$ is the matrix in single-particle states of the free time evolution operator,

$$u_0(t,t')_{ss'} = \langle s|\hat{u}_0(t,t')|s'\rangle. \tag{44}$$

2.3 Wick's theorem

Wick's theorem constitutes the basis of the diagrammatic expansion of Green functions. It requires that the initial density matrix has the uncorrelated form

$$\hat{\rho}_0 = \frac{e^{\hat{F}}}{\mathrm{Tr}\,e^{\hat{F}}} \tag{45}$$

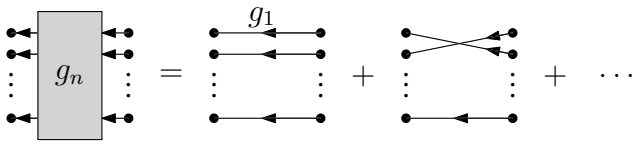


Fig. 3: *Diagrammatic illustration of Wick's theorem.*

with some single-particle operator

$$\hat{F} = \sum_{s',s} F_{s's} \hat{a}^\dagger_{s'} \hat{a}_s. \tag{46}$$

Let q denote the single-particle eigenstates of $\hat{F}$. In occupation number representation $\hat{\rho}_0$ has then the form

$$\hat{\rho}_0 = \sum_{n_{q_1}, n_{q_2}, \ldots} p_{n_{q_1}, n_{q_2}, \ldots} \left| n_{q_1}, n_{q_2}, \ldots \right\rangle \left\langle n_{q_1}, n_{q_2}, \ldots \right|, \tag{47}$$

with

$$p_{n_{q_1}, n_{q_2}, \ldots} = \frac{1}{\operatorname{Tr} e^{\hat{F}}} \prod_q \left(e^{F_q} \right)^{n_q}. \tag{48}$$

This shows that such a $\hat{\rho}_0$ describes a situation in which the particles occupy the states q completely independent of one another.
Wick theorem states [3] that

$$g(x_1, \ldots, x_n | x'_1, \ldots, x'_n) = \sum_{\substack{\text{permutations } P \\ \text{of } (1,\ldots,n)}} \zeta^P g(x_{P_1} | x'_1) \ldots g(x_{P_n} | x'_n) \tag{49}$$

if (and only if) $\hat{\rho}_0$ is of that specific form. The free n-particle Green function is the sum of all possible products of free single-particle Green functions which connect the external indices, compare Fig. 3
We introduce a more compact notation. We set

$$\hat{a}_x = \hat{a}^j_s(t)_0, \qquad \hat{a}^\dagger_{x'} = \hat{a}^{\dagger j'}_{s'}(t')_0. \tag{50}$$

Furthermore, for a permutation P of $(1, \ldots, n)$ we define the pairing $\hat{\mathcal{P}}_P$ of a product of n annihilators and n creators in the interactions picture as

$$\hat{\mathcal{P}}_P (\hat{a}_{x_1} \ldots \hat{a}_{x_n} \hat{a}^\dagger_{x'_n} \ldots \hat{a}^\dagger_{x'_1}) = i^n \zeta^P g(x_{P_1} | x'_1) \ldots g(x_{P_n} | x'_n). \tag{51}$$

The sum of all pairings is called total pairing $\hat{\mathcal{P}}$,

$$\hat{\mathcal{P}} = \sum_{\text{perm. } P} \hat{\mathcal{P}}_P. \tag{52}$$

Wick's theorem now reads

$$\operatorname{Tr} \hat{\mathcal{T}}_{\mathrm{c}}\, \hat{a}_{x_1} \ldots \hat{a}_{x_n} \hat{a}^\dagger_{x'_n} \ldots \hat{a}^\dagger_{x'_1} \rho_0 = \hat{\mathcal{P}} (\hat{a}_{x_1} \ldots \hat{a}_{x_n} \hat{a}^\dagger_{x'_n} \ldots \hat{a}^\dagger_{x'_1}). \tag{53}$$

We extend the definition of $\hat{\mathcal{P}}$ to arbitrary but particle number conserving products of generators and annihilators by symmetry under exchange of adjacent operators $\hat{a}$ or $\hat{a}^\dagger$ for Bosons and anti-symmetry for Fermions. Since $\hat{\mathcal{T}}_\mathrm{c}$ behaves identically we obtain quite generally

$$\mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,\hat{\tilde{a}}_{x_1}\ldots\hat{\tilde{a}}_{x_n}\hat{\rho}_0 = \hat{\mathcal{P}}(\hat{\tilde{a}}_{x_1}\ldots\hat{\tilde{a}}_{x_n}), \tag{54}$$

with $\hat{\tilde{a}}_x = \hat{a}_x$ or $\hat{a}_x^\dagger$ and an equal number of annihilators and creators in total. If the number of creators and annihilators is unequal, the left hand side vanishes as $\hat{H}_0$ and $\hat{\rho}_0$ commute with the operator of total particle number.

2.4 Diagrammatic expansion of the interacting Green function

It is rarely possible to compute Green functions of an interacting quantum system exactly. One approximation strategy is perturbation theory in powers of the interaction. It is convenient to express the numerous contributions to that expansion by diagrams. The derivation of the corresponding diagram rules is based on Wick's theorem. The core ideas applied here are identical to other Green function techniques like Matsubara formalism which is useful for systems in thermal equilibrium, or like zero-temperature formalism which describes systems in the ground state. Compare the diagram rules for instance to [6] or to lecture A4 by C. Friedrich.
We start by considering the time evolution operator in the interaction picture

$$\hat{u}_\mathrm{I}(t,t') = \hat{u}_0(t_0,t)\hat{u}(t,t')\hat{u}_0(t',t_0). \tag{55}$$

It satisfies the differential equation

$$i\partial_t\hat{u}_\mathrm{I}(t,t') = \hat{V}(t)_0\hat{u}_\mathrm{I}(t,t') \tag{56}$$

with solution

$$\hat{u}_\mathrm{I}(t,t') = \begin{cases} \hat{\mathcal{T}}\exp\left[-i\int_{t'}^{t}V(\tau)_0 d\tau\right], & t>t', \\ \hat{\tilde{\mathcal{T}}}\exp\left[i\int_{t}^{t'}V(\tau)_0 d\tau\right], & t<t'. \end{cases} \tag{57}$$

The expectation value of an operator can now be written

$$\langle\hat{A}\rangle(t) = \mathrm{Tr}\,\hat{u}(t_0,t)\hat{A}(t)\hat{u}(t,t_0)\hat{\rho}_0 \tag{58}$$

$$= \mathrm{Tr}\,\hat{u}_\mathrm{I}(t_0,t)\hat{A}(t)_0\hat{u}_\mathrm{I}(t,t_0)\hat{\rho}_0 \tag{59}$$

$$= \mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,e^{-i\int_{t_0}^{\infty}[\hat{V}^\mathrm{f}(\tau)_0-\hat{V}^\mathrm{b}(\tau)_0]d\tau}\hat{A}^\mathrm{f}(t)_0\hat{\rho}_0, \tag{60}$$

where we applied the same steps as in section 2.2 for the Heisenberg picture. For the Green function we get

$$G(x_1,\ldots,x_n|x_1',\ldots,x_n')$$

$$= (-i)^n\,\mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,e^{-i\int_{t_0}^{\infty}[\hat{V}^\mathrm{f}(\tau)_0-\hat{V}^\mathrm{b}(\tau)_0]d\tau}\hat{a}_{x_1}\ldots\hat{a}_{x_n}\hat{a}^\dagger_{x_n'}\ldots\hat{a}^\dagger_{x_1'}\hat{\rho}_0 \tag{61}$$

$$= (-i)^n\sum_{k=0}^{\infty}\frac{(-i)^k}{k!}\sum_{i_1,\ldots,i_k=\mathrm{f,b}}\int_{t_0}^{\infty}d\tau_1\ldots d\tau_k\,\mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,\mathrm{sgn}(i_1)\hat{V}^{i_1}(\tau_1)_0\ldots\mathrm{sgn}(i_k)\hat{V}^{i_k}(\tau_k)_0$$
$$\times\,\hat{a}_{x_1}\ldots\hat{a}_{x_n}\hat{a}^\dagger_{x_n'}\ldots\hat{a}^\dagger_{x_1'}\hat{\rho}_0, \tag{62}$$

where

$$\operatorname{sgn} i = \delta_{i,\mathrm{f}} - \delta_{i,\mathrm{b}}. \tag{63}$$

After inserting the second quantized expression for $\hat{V}$ we can apply Wick's theorem Eq. (54). Let us set

$$\bar{\nu}_{x_1' x_2' x_1 x_2} = \bar{\nu}^{j_1' j_2' j_1 j_2}_{s_1' s_2' s_1 s_2}(t_1', t_2', t_1, t_2) \tag{64}$$

$$= \delta(t_1' - t_1)\delta(t_2' - t_1)\delta(t_2 - t_1)\delta_{j_1' j_1}\delta_{j_2' j_1}\delta_{j_2 j_1} \operatorname{sgn}(j_1)\bar{v}_{s_1' s_2' s_1 s_2}. \tag{65}$$

Then we find

$$G(x_1, \ldots, x_n | x_1', \ldots, x_n') = (-i)^n \sum_{k=0}^{\infty} \frac{(-i)^k}{k!} \frac{\bar{\nu}_{y_{11}' y_{12}' y_{11} y_{12}}}{4} \ldots \frac{\bar{\nu}_{y_{k1}' y_{k2}' y_{k1} y_{k2}}}{4}$$
$$\times \hat{\mathcal{P}}(\hat{a}^\dagger_{y_{11}'} \hat{a}^\dagger_{y_{12}'} \hat{a}_{y_{12}} \hat{a}_{y_{11}} \ldots \hat{a}^\dagger_{y_{k1}'} \hat{a}^\dagger_{y_{k2}'} \hat{a}_{y_{k2}} \hat{a}_{y_{k1}} \hat{a}_{x_1} \ldots \hat{a}_{x_n} \hat{a}^\dagger_{x_n'} \ldots \hat{a}^\dagger_{x_1'}). \tag{66}$$

Here, all multi-indices appearing twice are contracted, which means that contour indices and state indices are summed over and times are integrated over from t_0 to ∞. In abbreviated notation this reads

$$G(x_1, \ldots, x_n | x_1', \ldots, x_n') = (-i)^n \sum_{k=0}^{\infty} \frac{(-i)^k}{k!} \frac{\bar{\nu}_1}{4} \ldots \frac{\bar{\nu}_k}{4} \hat{\mathcal{P}}(\hat{A}_1 \ldots \hat{A}_k \hat{a}_{x_1} \ldots \hat{a}^\dagger_{x_1'}), \tag{67}$$

with

$$\bar{\nu}_i = \bar{\nu}_{y_{i1}' y_{i2}' y_{i1} y_{i2}}, \tag{68}$$

$$\hat{A}_i = \hat{a}^\dagger_{y_{i1}'} \hat{a}^\dagger_{y_{i2}'} \hat{a}_{y_{i2}} \hat{a}_{y_{i1}}. \tag{69}$$

The pairings that contribute to the total pairing in this expression for the Green function will be expressed as diagrams.
However, only certain pairings need to be considered, namely those which cannot be factorized in the form

$$\hat{\mathcal{P}}_P(\hat{A}_1 \ldots \hat{A}_k \hat{a}_{x_1} \ldots \hat{a}^\dagger_{x_1'}) = \hat{\mathcal{P}}_{P_1}(\hat{A}_{i_1} \ldots \hat{A}_{i_l}) \hat{\mathcal{P}}_{P_2}(\hat{A}_{i_{l+1}} \ldots \hat{A}_{i_k} \hat{a}_{x_1} \ldots \hat{a}^\dagger_{x_1'}) \tag{70}$$

for some l in $\{1, \ldots, k\}$. Below, these pairings correspond to diagrams in which each vertex is connected to at least one external point via lines and other vertices. We call the corresponding permutation "connected to the external fields" and denote the sum of all such pairings the total pairing connected to the external fields $\hat{\mathcal{P}}_{\mathrm{ce}}$. From

$$\hat{\mathcal{P}}(\hat{A}^k \hat{a}_{x_1} \ldots \hat{a}^\dagger_{x_1'}) = \sum_{l=0}^{k} \binom{k}{l} \hat{\mathcal{P}}(\hat{A}^l) \hat{\mathcal{P}}_{\mathrm{ce}}(\hat{A}^{k-l} \hat{a}_{x_1} \ldots \hat{a}^\dagger_{x_1'}). \tag{71}$$

we find

$$G(x_1, \ldots, x_n | x_1', \ldots, x_n') = (-i)^n \sum_{l=0}^{\infty} \frac{(-i)^l}{l!} \left(\frac{\bar{\nu}}{4}\right)^l \hat{\mathcal{P}}(\hat{A}^l)$$
$$\times \sum_{m=0}^{\infty} \frac{(-i)^m}{m!} \left(\frac{\bar{\nu}}{4}\right)^m \hat{\mathcal{P}}_{\mathrm{ce}}(\hat{A}^m \hat{a}_{x_1} \ldots \hat{a}^\dagger_{x_1'}). \tag{72}$$

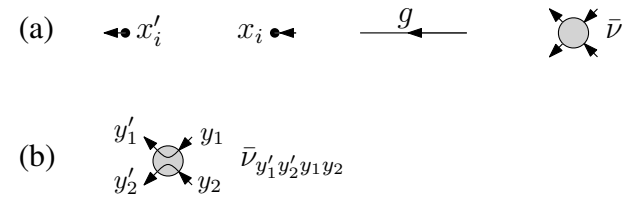


Fig. 4: *(a) The constituents of diagrams for the expansion of Green functions in powers of the interaction are starting points x'_i and end points x_i (external vertices), lines representing free single-particle Green functions g, and vertices representing interaction amplitudes $\bar{\nu}$. (b) A vertex, after multi-indices have been assigned to its connectors.*

Speaking diagrammatically, the first series represents the sum of all diagrams without any external vertices (the "vacuum diagrams"). It has the value

$$\sum_{l=0}^{\infty} \frac{(-i)^l}{l!} \left(\frac{\bar{\nu}}{4}\right)^l \hat{\mathcal{P}}(\hat{A}^l) = \mathrm{Tr}\, \hat{\mathcal{T}}_{\mathrm{c}}\, e^{-i\int_{t_0}^{\infty}[\hat{V}^{\mathrm{f}}(\tau)_0 - \hat{V}^{\mathrm{b}}(\tau)_0]d\tau} \hat{\rho}_0 \tag{73}$$

$$= \mathrm{Tr}\, u_{\mathrm{I}}(t_0,\infty)\hat{u}_{\mathrm{I}}(\infty,t_0)\hat{\rho}_0 \tag{74}$$

$$= 1. \tag{75}$$

Hence

$$G(x_1,\ldots,x_n|x'_1,\ldots,x'_n) = (-i)^n \sum_{m=0}^{\infty} \frac{(-i)^m}{m!} \left(\frac{\bar{\nu}}{4}\right)^m \hat{\mathcal{P}}_{\mathrm{ce}}\left(\hat{A}^m \hat{a}_{x_1} \ldots \hat{a}^{\dagger}_{x'_1}\right) \tag{76}$$

$$= \sum_{m=0}^{\infty} \frac{i^m}{m!} \underbrace{\frac{\bar{\nu}}{4} \ldots \frac{\bar{\nu}}{4}}_{m \text{ times}} \sum_{\substack{P \text{ conn. to} \\ \text{ext. fields}}} \zeta^P \underbrace{g \ldots g}_{(2m+n) \text{ times}} . \tag{77}$$

where all internal indices are contracted. The different permutations that are summed over can be represented by diagrams composed of directed lines that connect vertices. A directed line corresponds to a free Green function g and a vertex corresponds to an interaction amplitude $\bar{\nu}$. External vertices are start and end points of lines, as depicted in Fig. 4 (a). Due to

$$\bar{\nu}_{y'_1 y'_2 y_1 y_2} = \zeta \bar{\nu}_{y'_2 y'_1 y_1 y_2} = \zeta \bar{\nu}_{y'_1 y'_2 y_2 y_1}, \tag{78}$$

two permutations yield the same value, if they differ only by interchange of the two incoming indices y_1, y_2 of a vertex or by interchange of the two outgoing indices y'_1, y'_2 of a vertex. The $2 \times 2 = 4$ identical permutations can be represented by a single vertex without indices and which counts fourfold. However, counting vertices fourfold is double-counting, if two lines connect the same vertices in the same direction. We correct these cases by dividing by two. Permutations that differ only by interchange of the complete sets of indices of one vertex with another vertex yield the same value. A single diagram with unlabeled vertices represents all these permutations. For a diagram of order V^m there are $m!$ such permutations, except when the diagram has a certain symmetry. Finally it turns out that ζ^P can be determined from the number of closed loops in the diagram and from the order, in which the diagram connects the start and end points. These considerations are taken into account by the following diagram rules.

Rules for calculating the contribution of order V^m to the Green-function $G(x_1,\ldots,x_n|x'_1,\ldots,x'_n)$, using Hugenholtz diagrams. Draw all distinct diagrams composed of n external

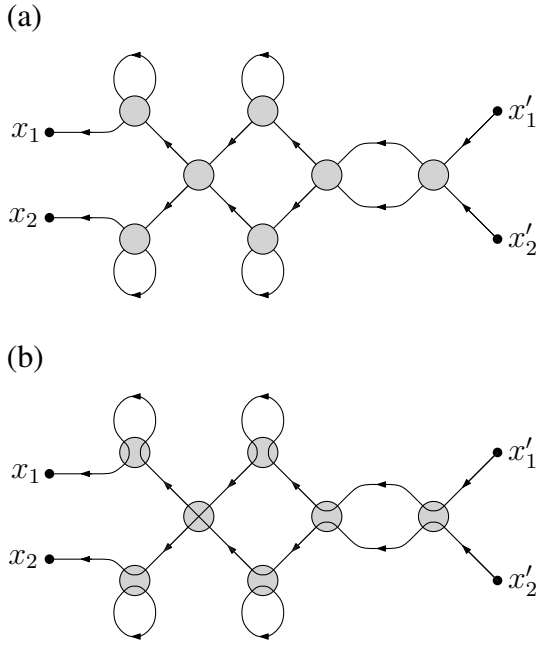


Fig. 5: *(a) An example diagram of order $m = 7$ contributing to the two-particle Green function. The number of pairs of equivalent lines is $n_{\text{eq}} = 1$, and the symmetry factor is $S = 2$. (b) After having chosen arbitrarily how lines continue inside the vertices, we find $n_{\text{loop}} = 2$ and $\zeta^P = \zeta$. Other choices can yields different values for n_{loop} and ζ^P. A possible change in $\zeta^{n_{\text{loop}}}\zeta^P$ is counterbalanced by different factors ζ from the vertices, compare Eq. (78).*

starting points x'_i , n external end points x_i, and m vertices that are connected by directed lines. Two diagrams are distinct, if they cannot be deformed so as to coincide completely including the direction of arrows and the external labels. Consider only diagrams in which each vertex is connected (via lines and other vertices) to at least one of the external points. The contribution of order V^m is the sum of all such diagrams, where a single diagram is evaluated as follows.
Let n_{eq} count the pairs of equivalent lines in the diagram. Two lines are called equivalent if they connect the same vertices and run in the same direction. Assign labels $k = 1, \ldots, m$ to the m vertices. Determine the symmetry factor S as the number of permutations of these labels that map the diagrams onto itself. (Mind to count the identity permutation.) To the incoming connectors of each vertex assign multi-indices $y_{k,1}, y_{k,2}$ (in arbitrary order), and to the outgoing connectors assign the indices $y'_{k,1}, y'_{k,2}$. The k-th vertex then has the value $\bar{\nu}_{y'_{k,1}y'_{k,2}y_{k,1}y_{k,2}}$. The directed line starting at some (internal or external) index y' and ending at some (internal or external) index y has the value $g(y|y')$. Define n_{loop} to be the number of loops made up by the lines. [At the vertices, one line continues from $y_{k,1}$ to $y'_{k,1}$ and the other from $y_{k,2}$ to $y'_{k,2}$, compare Fig. 4 (b).] Finally, the starting point x'_i being connected via lines to the end point x_{P_i} determines a permutation P. [This is a permutation of $\{1, \ldots, n\}$ and different from the permutation appearing in Eq. (77)]. The value of the diagram is then

$$\frac{\zeta^P \zeta^{n_{\text{loop}}}}{2^{n_{\text{eq}}} S} \left(\prod i\bar{\nu}\right) \left(\prod g\right). \tag{79}$$

Here, all internal multi-indices are contracted which means summation over the contour index and single-particle state and integration over time from t_0 to ∞. The prefactor of an example diagram is discussed in Figure 5. Note that for each vertex there remains only one independent

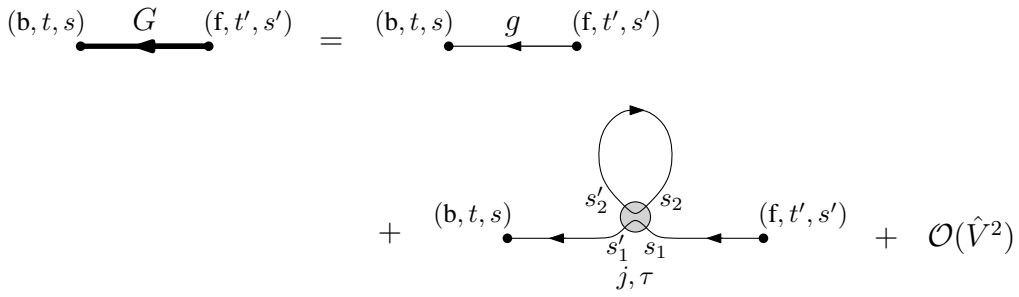


Fig. 6: *Expansion of $G^{>}_{s|s'}(t|t')$ up to first order in the interaction.*

time integration and one independent contour-index sum, since

$$\bar{\nu}_{y_1' y_2' y_1 y_2} \sim \delta(t_1' - t_1)\delta(t_2' - t_1)\delta(t_2 - t_1)\delta_{j_1' j_1}\delta_{j_2' j_1}\delta_{j_2 j_1}. \tag{80}$$

As an example we discuss first order perturbation theory for $G^{>}_{s|s'}(t|t')$. The corresponding diagrams are depicted in Fig. 6. They yield

$$\begin{aligned} &G^{>}_{s|s'}(t|t') \\ &= g^{>}_{s|s'}(t|t') + \frac{\zeta^0\zeta^1}{2^0 \cdot 1}\sum_{j=\mathrm{f,b}}\int_{t_0}^{\infty} d\tau\, g^{\mathrm{b}|j}_{s|s_1'}(t|\tau)\left[i\,\mathrm{sgn}(j)\bar{v}_{s_1' s_2' s_1 s_2}(\tau)\right] g^{j|j}_{s_2|s_2'}(\tau|\tau) g^{j|\mathrm{f}}_{s_1|s'}(\tau|t') + \mathcal{O}(\hat{V}^2) \qquad (81) \\ &= g^{>}_{s|s'}(t|t') + \int_{t_0}^{\infty} d\tau\, \left[g^{>}_{s|s_1'}(t|\tau) g^{\mathrm{c}}_{s_1|s'}(\tau|t') - g^{\tilde{\mathrm{c}}}_{s|s_1'}(t|\tau) g^{>}_{s_1|s'}(\tau|t')\right] \qquad (82) \\ &\qquad\qquad \times \bar{v}_{s_1' s_2' s_1 s_2}(\tau)\left[i\zeta g^{<}_{s_2|s_2'}(\tau|\tau)\right] + \mathcal{O}(\hat{V}^2). \qquad (83) \end{aligned}$$

Here, $i\zeta g^{<}_{s_2|s_2'}(\tau|\tau) = \mathrm{Tr}\left(\hat{a}^{\dagger}_{s_2'}\hat{a}_{s_2}\right)(\tau)_0\rho_0$ is the occupation matrix of the noninteracting system at time τ. The particles interact via $\bar{v}$ with the extra-particle whose propagation is described by $G^{>}$. The two addends in square brackets in Eq. (82) describe that this Hartree-Fock type interaction takes place either on the forward or on the backward contour.

2.5 Self-energy

The concept of a self-energy exists in Keldysh formalism in the same way as in other Green function formalisms, compare for instance [6] and lecture A4 by C. Friedrich. A peculiarity of Keldysh formalism is that the self-energy bears two contour indices which make it a four component object.

Consider all diagrams contributing to the single-particle Green function $G(x|x')$. In each such diagram, the starting point x' is connected to the end point x by a sequence of lines and other vertices. This is a consequence of particle conservation in our model. Choose a diagram and remove one specific line. If this removal disconnects x from x', we call the line a branch line of the diagram.

Let us classify the diagrams according to their number of branch lines. There is no diagram containing no branch line. There is a single diagram containing a single branch line, namely the diagram representing the free Green function, compare Fig. 7 (a). The diagrams with two branch

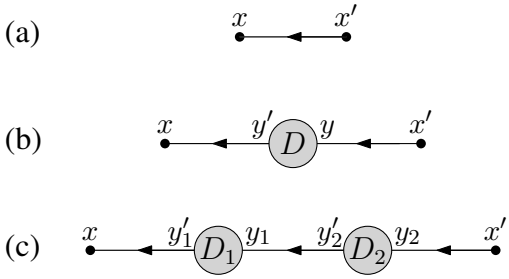


Fig. 7: *Diagrams contributing to the single-particle Green function $G(x|x')$ classified by their number of branch lines. (a) One branch line. (b) Two branch lines. The sub-diagram D is one-particle irreducible. (c) Three branch lines. D_1 and D_2 are one-particle irreducible.*

lines have the structure depicted in Fig. 7 (b). The sub-diagram D is a so called amputated one-particle irreducible one-particle diagram. It is an amputated diagram, since the lines connecting it with the external points x and x' are not part of the diagram. Hence, $D = D(y'|y)$ is a function of the multi-indices y', y which are to be contracted with the two remaining free Green functions $g(x|y')$ and $g(y|x')$. D is called one-particle irreducible, since it does not contain any branch line. This means that removing any single line from that diagram does not disconnect it. When we evaluate such an amputated one-particle irreducible one-particle diagram as

$$\frac{\zeta^P \zeta^{n_{\text{loop}}}}{2^{n_{\text{eq}}} S} \left(\prod i\bar{\nu}\right) \left(\prod g\right), \tag{84}$$

then the product $\prod g$ does no longer comprise the lines $g(x|y')$ and $g(y|x')$, and the implicit summation over the internal multi-indices does not include the indices y', y. The total diagram in Fig. 7 (b) has the value

$$g(x|y')D(y'|y)g(y|x'), \tag{85}$$

where y and y' are summed over.
The diagrams with three branch lines have the structure indicated in Fig. 7 (c). The value of such a diagram is

$$g(x|y'_1)D_1(y'_1|y_1)g(y_1|y'_2)D_2(y'_2|y_2)g(y_2|x'). \tag{86}$$

This equation is based on the fact that there are no equivalent lines connecting a vertex from D_1 with a vertex from D_2 and that interchanging a vertex label from D_1 with one of D_2 is no symmetry of the diagram. Additionally, each loop is either in D_1 or in D_2. A diagram with k branch lines is given by

$$g(x|y'_1)D_1(y'_1|y_1)g(y_1|y'_2)\ldots D_{k-1}(y'_{k-1}|y_{k-1})g(y_{k-1}|x'). \tag{87}$$

The self-energy $\Sigma(y'|y)$ is defined to be the sum of all amputated one-particle irreducible one-particle diagrams $D(y'|y)$. The sum of all diagrams with the structure given in Fig. 7 (b) is then

$$g(x|y')\Sigma(y'|y)g(y|x'), \tag{88}$$

and the sum of all diagrams with the structure from Fig. 7 (c) is

$$g(x|y'_1)\Sigma(y'_1|y_1)g(y_1|y'_2)\Sigma(y'_2|y_2)g(y_2|x'). \tag{89}$$

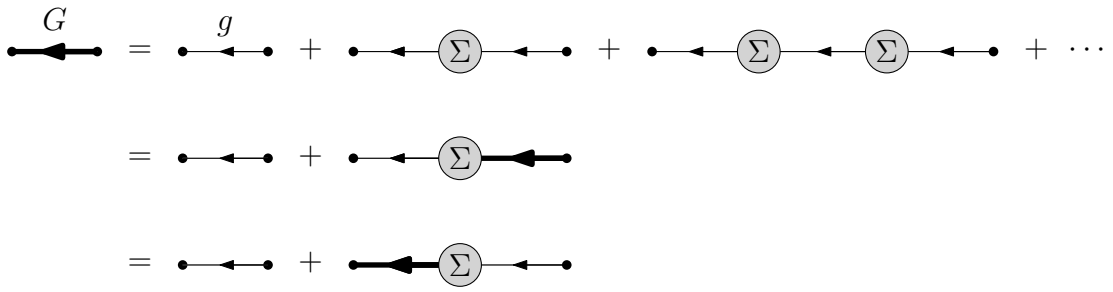


Fig. 8: *Dyson's equation*

Summation over the number of branch lines leads to Dyson's equation for the single-particle Green function,

$$G = g + g\Sigma g + g\Sigma g\Sigma g + \ldots \tag{90}$$
$$= g + g\Sigma G \tag{91}$$
$$= g + G\Sigma g, \tag{92}$$

which is represented in Fig. 8. Here the product is a short notation for the contraction of indices. In particular, the self-energy has two contour indices, and the components can be arranged in a matrix as

$$\Sigma^{j'|j} = \begin{pmatrix} \Sigma^{\mathrm{f|f}} & \Sigma^{\mathrm{f|b}} \\ \Sigma^{\mathrm{b|f}} & \Sigma^{\mathrm{b|b}} \end{pmatrix}_{j'j} = \begin{pmatrix} \Sigma^{\mathrm{c}} & \Sigma^{<} \\ \Sigma^{>} & \Sigma^{\tilde{\mathrm{c}}} \end{pmatrix}_{j'j}. \tag{93}$$

The solution of Dyson's equation is

$$G = (g^{-1} - \Sigma)^{-1}. \tag{94}$$

Hence, the self-energy allows to compute the single-particle Green function, which in turn describes all single-particle features of the system. Due to the series structure of Dyson's equation, a finite order approximation to the self-energy results in an approximation for the Green function including infinitely many diagrams.

2.6 Keldysh rotation

The Keldysh rotation [7] is a unitary transformation in the contour indices given by

$$\hat{a}^1 = \frac{\hat{a}^{\mathrm{f}} - \hat{a}^{\mathrm{b}}}{\sqrt{2}}, \tag{95}$$
$$\hat{a}^2 = \frac{\hat{a}^{\mathrm{f}} + \hat{a}^{\mathrm{b}}}{\sqrt{2}}. \tag{96}$$

We call $l = 1, 2$ the Keldysh index as opposed to the contour index $j = \mathrm{f}, \mathrm{b}$. Let $z = (l, t, s)$ be the corresponding multi-index, and define the rotation matrix

$$K_{z_1 x_2} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}_{l_1 j_2} \delta(t_1 - t_2)\delta_{s_1 s_2}, \tag{97}$$
$$K^\dagger = K^{-1}. \tag{98}$$

Then

$$\hat{a}_z = K_{zx}\hat{a}_x, \tag{99}$$
$$\hat{a}_z^\dagger = \hat{a}_x^\dagger K_{xz}^\dagger = \hat{a}_x^\dagger K_{xz}^{-1}, \tag{100}$$

with summation convention for repeated multi-indices. The Keldysh rotated n-particle Green functions is given by

$$G(z_1,\ldots,z_n|z_1',\ldots,z_n') = (-i)^n \operatorname{Tr} \hat{\mathcal{T}}_\mathrm{c}\, \hat{a}_{s_1}^{l_1}(t_1)_\mathrm{H} \ldots \hat{a}_{s_n}^{l_n}(t_n)_\mathrm{H} \hat{a}_{s_n'}^{\dagger\, l_n'}(t_n')_\mathrm{H} \ldots \hat{a}_{s_1'}^{\dagger\, l_1'}(t_1')_\mathrm{H}\hat{\rho}_0 \tag{101}$$
$$= K_{z_1x_1}\ldots K_{z_nx_n} G(x_1,\ldots,x_n|x_1',\ldots,x_n') K^{-1}_{x_1'z_1'}\ldots K^{-1}_{x_n'z_n'}. \tag{102}$$

We transform vertices and the self-energy accordingly,

$$\bar{\nu}_{z_1'z_2'z_1z_2} = K_{z_1'x_1'}K_{z_2'x_2'}\bar{\nu}_{x_1'x_2'x_1x_2}K^{-1}_{x_1z_1}K^{-1}_{x_2z_2}, \tag{103}$$
$$\Sigma(z'|z) = K_{z'x'}\Sigma(x'|x)K^{-1}_{xz}. \tag{104}$$

Therefore index contractions of vertices and (free) Green functions remain unchanged and all diagram rules formulated above can be applied as well for Keldysh rotated Green functions.
The nontrivial components of the Keldysh rotated single-particle Green function, are called retarded , advanced, Keldysh Green function,

$$\begin{pmatrix} G^{1|1} & G^{1|2} \\ G^{2|1} & G^{2|2} \end{pmatrix} = \begin{pmatrix} G^\mathrm{P} & G^\mathrm{A} \\ G^\mathrm{R} & G^\mathrm{K}+G^\mathrm{P} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} G^\mathrm{c} & G^< \\ G^> & G^{\tilde{\mathrm{c}}} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{105}$$

They are given by

$$G^\mathrm{R}(t|t') = \frac{1}{2}\left[G^\mathrm{c} - G^< + G^> - G^{\tilde{\mathrm{c}}}\right](t|t') = \Theta(t-t')\left[G^>(t|t') - G^<(t|t')\right], \tag{106}$$
$$G^\mathrm{A}(t|t') = \frac{1}{2}\left[G^\mathrm{c} + G^< - G^> - G^{\tilde{\mathrm{c}}}\right](t|t') = \Theta(t'-t)\left[G^<(t|t') - G^>(t|t')\right], \tag{107}$$
$$G^\mathrm{K} = G^< + G^>. \tag{108}$$

with $\Theta(0) = 1/2$ (in particular $G^\mathrm{R}(t|t) = -i/2$). Without recourse to contour ordering they can be written

$$G^\mathrm{R}_{s|s'}(t|t') = -i\Theta(t-t')\operatorname{Tr}\left[\hat{a}_s(t)_\mathrm{H}, \hat{a}_{s'}^\dagger(t')_\mathrm{H}\right]_{-\zeta}\hat{\rho}_0, \tag{109}$$
$$G^\mathrm{A}_{s|s'}(t|t') = i\Theta(t'-t)\operatorname{Tr}\left[\hat{a}_s(t)_\mathrm{H}, \hat{a}_{s'}^\dagger(t')_\mathrm{H}\right]_{-\zeta}\hat{\rho}_0 = G^\mathrm{R}_{s'|s}(t'|t)^*, \tag{110}$$
$$G^\mathrm{K}_{s|s'}(t|t') = -i\operatorname{Tr}\left[\hat{a}_s(t)_\mathrm{H}, \hat{a}_{s'}^\dagger(t')_\mathrm{H}\right]_{\zeta}\hat{\rho}_0 = -G^\mathrm{K}_{s'|s}(t'|t)^*. \tag{111}$$

The $(1|1)$ component vanishes except for a small "patch",

$$G^\mathrm{P}(t|t') = \frac{1}{2}\left[G^\mathrm{c} - G^< - G^> + G^{\tilde{\mathrm{c}}}\right](t|t') = \left\{\begin{matrix} 0, & t\neq t' \\ \frac{i}{2}, & t=t' \end{matrix}\right\} = g^\mathrm{P}(t|t'). \tag{112}$$

Since g^P vanishes except at a single point, it does not contribute when its time arguments are integrated over. The only relevance of the patch Green function is for lines that start and end at the same vertex, as does the loop line in the first order diagram of Fig. 6. In such cases, on encounters the expression

$$\frac{1}{2}\left[g^\mathrm{K}(t|t) + 2g^\mathrm{P}(t,t)\right] = g^<(t|t), \tag{113}$$

yielding a Hartree-Fock like contribution as described after Eq. (83).
The fact that G^{P} essentially vanishes simplifies computations and is one of the motivations for the Keldysh rotation. As $G^{\mathrm{A}} = G^{\mathrm{R}\,\dagger}$, and $G^{\mathrm{P}} = g^{\mathrm{P}}$, all single particle functions can be derived from the two functions G^{R} and G^{K} alone. Retarded Green functions are known from linear response theory. Their inbuilt causal structure of the cause preceeding the effect becomes visible in the factor $\Theta(t-t')$ in Eq. (109). This factor is directly connected to the analytic behaviour of the retarded Green function in frequency representation, as described in the following section. More generally, multi-particle Green functions vanish, if the Keldysh index associated with the greatest time argument equals 1 [10]. In thermal equilibrium, G^{R} is directly connected to the Matsubara Green function, as discussed in section 2.8.
The explicit results for the components of the free single-particle Green function are

$$g^{\mathrm{R}}(t,t') = -i\Theta(t-t')u_0(t,t') = g^{\mathrm{A}}(t',t)^{\dagger} \tag{114}$$
$$g^{\mathrm{K}}(t,t') = -iu_0(t,t_0)(1+\zeta 2n)u_0(t_0,t') \tag{115}$$
$$= -ig^{\mathrm{R}}(t,t_0)(1+\zeta 2n)g^{\mathrm{A}}(t_0,t'). \tag{116}$$
$$= \left(g^{\mathrm{R}}\Sigma^{\mathrm{K}}_{\mathrm{ini}}g^{\mathrm{A}}\right)(t,t'). \tag{117}$$

In the last line, multiplication includes summation over single-particle states and integration over time. We defined the initial Keldysh self-energy

$$\Sigma^{\mathrm{K}}_{\mathrm{ini}}(t'|t) = -i(1+\zeta 2n)\delta(t'-t)\delta(t-t_0). \tag{118}$$

The inverse Keldysh rotated Green function is

$$G^{-1} = \begin{pmatrix} G^{\mathrm{P}} & G^{\mathrm{A}} \\ G^{\mathrm{R}} & G^{\mathrm{K}}+G^{\mathrm{P}} \end{pmatrix}^{-1} = \begin{pmatrix} -G^{\mathrm{R}\,-1}G^{\mathrm{K}}G^{\mathrm{A}\,-1} & G^{\mathrm{R}\,-1} \\ G^{\mathrm{A}\,-1} & 0 \end{pmatrix}, \tag{119}$$
$$g^{-1} = \begin{pmatrix} -\Sigma^{\mathrm{K}}_{\mathrm{ini}} & g^{\mathrm{R}\,-1} \\ g^{\mathrm{A}\,-1} & 0 \end{pmatrix}. \tag{120}$$

When checking the inverse, note that $G^{\mathrm{P}}G^{\mathrm{R}\,-1} = 0$ since G^{P} vanishes except at one point and the internal time is integrated over. We conclude for the Keldysh-rotated self-energy

$$g^{-1} - G^{-1} = \Sigma = \begin{pmatrix} \Sigma^{1|1} & \Sigma^{1|2} \\ \Sigma^{2|1} & \Sigma^{2|2} \end{pmatrix} = \begin{pmatrix} \Sigma^{\mathrm{K}} & \Sigma^{\mathrm{R}} \\ \Sigma^{\mathrm{A}} & 0 \end{pmatrix}. \tag{121}$$

It can be shown that the self-energy (and more generally, any one-particle irreducible multi-particle vertex function) vanishes, if the Keldysh index associated with the greatest time argument equals 2 [10]. In particular $\Sigma^{\mathrm{R}}(t'|t) \sim \Theta(t'-t)$ and $\Sigma^{2|2} = 0$. The components of Dyson's equation read

$$G^{\mathrm{R}} = \left(g^{\mathrm{R}\,-1} - \Sigma^{\mathrm{R}}\right)^{-1}, \tag{122}$$
$$G^{\mathrm{A}} = \left(g^{\mathrm{A}\,-1} - \Sigma^{\mathrm{A}}\right)^{-1}, \tag{123}$$
$$G^{\mathrm{K}} = G^{\mathrm{R}}\left(\Sigma^{\mathrm{K}} + \Sigma^{\mathrm{K}}_{\mathrm{ini}}\right)G^{\mathrm{A}}. \tag{124}$$

In particular, there is a decoupled Dyson's equation for the retarded component alone. The structure of Dyson's equation in the Keldysh component is interesting as well. Information on the initial density matrix is stored in $\Sigma^{\mathrm{K}}_{\mathrm{ini}}$. This information is propagated from t_0 to the time

arguments $t, t' > t_0$ of G^{K} via G^{R} and G^{A}. If inelastic processes are present in the system then $G^{\mathrm{R,A}}$ will decay for long time spans, compare Eq. (128) below. Consequently, the information on the initial density matrix gets suppressed for increasing times. On the other hand, at all times later than t_0, the interaction influences the particle statistics in a way encoded in Σ^{K}. That information is propagated as well to the time arguments t, t' of G^{K} and may become the dominant contribution. Structurally identical equations exist for $G^{<}$ and $G^{>}$. $G^{<}$ is related to the statistics of particles, $G^{>}$ to that of holes and $G^{\mathrm{K}} = G^{<} + G^{>}$ to a combination of both.
Finally, the Keldysh rotated two-particle vertex is

$$\bar{\nu}_{z_1' z_2' z_1 z_2} = \begin{cases} \delta(t_1' - t_1)\delta(t_2' - t_1)\delta(t_2 - t_1)\frac{\bar{v}_{s_1' s_2' s_1 s_2}}{2}(t_1), & l_1' + l_2' + l_1 + l_2 \text{ odd}, \\ 0, & l_1' + l_2' + l_1 + l_2 \text{ even}. \end{cases} \tag{125}$$

For single-particle functions another convention for the Keldysh rotation is popular, in which the definition of $\hat{a}^1$ and $\hat{a}^2$ is interchanged (but not that of $\hat{a}^{\dagger\,1}$ and $\hat{a}^{\dagger\,2}$) [8]. In this convention, the single-particle Green function and the self-energy have the structure,

$$\begin{pmatrix} G^{\mathrm{R}} & G^{\mathrm{K}} + G^{\mathrm{P}} \\ G^{\mathrm{P}} & G^{\mathrm{A}} \end{pmatrix}, \qquad \begin{pmatrix} \Sigma^{\mathrm{R}} & \Sigma^{\mathrm{K}} \\ 0 & \Sigma^{\mathrm{A}} \end{pmatrix}. \tag{126}$$

2.7 Fourier transformation

In this section, we consider the case of a time independent Hamiltonian $H = H_0 + V$. If inelastic processes lead to a decay of long-term correlations it may happen that the system state becomes stationary for times much later than the initial time t_0. This means that the Green functions acquire time translational invariance in that limit,

$$\lim_{t_0 \to -\infty} G(t_1, \ldots, t_n | t_1', \ldots, t_n') = \lim_{t_0 \to -\infty} G(t_1 - t, \ldots, t_n - t | t_1' - t, \ldots, t_n' - t) \tag{127}$$

for all finite t. If one is only interested in properties of the stationary state, it is advantageous to work in Fourier space, where time translation invariance leads to frequency conservation. The decay of long-term correlations,

$$\lim_{|t_i - t_j| \to \infty} G(t_1, \ldots, t_n | t_1', \ldots, t_n') = 0, \quad \text{for all } t_i, t_j \in \{t_1, \ldots, t_n'\}, \tag{128}$$

assures the convergence of Fourier transformations in $2n - 1$ of the Green function's time arguments. The Fourier transformation in the remaining time argument leads to a δ-function that assures frequency conservation.
We set $t_0 = -\infty$ and define the Fourier transformed n-particle Green function by

$$G(\omega_1, \ldots, \omega_n | \omega_1', \ldots, \omega_n') = \int_{-\infty}^{\infty} e^{i(\omega_1 t_1 + \cdots + \omega_n t_n - \omega_1' t_1' - \cdots - \omega_n' t_n')} G(t_1, \ldots, t_n | t_1', \ldots, t_n') dt_1 \ldots dt_n'. \tag{129}$$

Let $\xi = (l, \omega, s)$ [or (j, ω, s)] as opposed to $z = (l, t, s)$ [or $x = (j, t, s)$] be the multi-index with frequency instead of time. Consider the transformation matrix

$$F_{\xi_1 z_2} = \delta_{l_1 l_2} e^{i\omega_1 t_2} \delta_{s_1 s_2}. \tag{130}$$

If we define contraction of frequencies with a prefactor $(2\pi)^{-1}$,

$$A_\omega B_\omega = \frac{1}{2\pi}\int d\omega A_\omega B_\omega, \tag{131}$$

then the identity in Fourier space is $1_{\omega\omega'} = 2\pi\delta(\omega-\omega')$, and F is unitary,

$$F^\dagger = F^{-1}. \tag{132}$$

The Fourier transformed Green function can now be written

$$G(\xi_1,\ldots,\xi_n|\xi_1',\ldots,\xi_n') = F_{\xi_1 z_1}\ldots F_{\xi_n z_n} G(z_1,\ldots,z_n|z_1',\ldots,z_n') F^{-1}_{z_1'\xi_1'}\ldots F^{-1}_{z_n'\xi_n'}, \tag{133}$$

with summation convention for repeated multi-indices. We transform vertices and the self-energy accordingly,

$$\bar\nu_{\xi_1'\xi_2'\xi_1\xi_2} = F_{\xi_1' z_1'} F_{\xi_2' z_2'} \bar\nu_{z_1' z_2' z_1 z_2} F^{-1}_{z_1\xi_1} F^{-1}_{z_2\xi_2}, \tag{134}$$

$$\Sigma(\xi'|\xi) = F_{\xi' z'}\Sigma(z'|z)F^{-1}_{z\xi}. \tag{135}$$

Therefore index contractions of vertices and (free) Green functions remain unchanged and all diagram rules formulated above can be applied as well for the Fourier transformed Green functions. Due to frequency conservation the number of independent frequency integrations to be performed for the evaluation of a diagram is strongly reduced.
From

$$G(t_1,\ldots,t_n|t_1',\ldots,t_n') = G(0,t_2-t_1,\ldots,t_n-t_1|t_1'-t_1,\ldots,t_n'-t_1) \tag{136}$$

follows that the Green function in Fourier space conserves frequency,

$$\begin{aligned} &G(\omega_1,\ldots,\omega_n|\omega_1',\ldots,\omega_n') \\ &= \int_{-\infty}^{\infty} e^{i(\omega_1 t_1+\cdots+\omega_n t_n-\omega_1' t_1'-\cdots-\omega_n' t_n')} G(0,t_2-t_1,\ldots,t_n-t_1|t_1'-t_1,\ldots,t_n'-t_1)dt_1\ldots dt_n' \end{aligned} \tag{137}$$

$$= 2\pi\delta(\omega_1+\cdots+\omega_n-\omega_1'-\cdots-\omega_n')G(t_1=0,\omega_2,\ldots,\omega_n|\omega_1',\ldots,\omega_n'). \tag{138}$$

Especially for the single particle Green function we set

$$G(\omega) := G(\omega|t'=0) = \int_{-\infty}^{\infty} e^{i\omega t}G(t|0)dt = \int_{-\infty}^{\infty} e^{-i\omega t'}G(0|t')dt' = G(t=0|\omega), \tag{139}$$

and find that

$$G^{\mathrm{R}}(z) = \int_0^\infty e^{izt}G^{\mathrm{R}}(t|0)dt \tag{140}$$

is analytic in the upper half plane of z, since

$$e^{izt} \sim e^{-\operatorname{Im}(z)t} \qquad \text{and} \qquad G^{\mathrm{R}}(t|0) \xrightarrow{t\to\infty} 0. \tag{141}$$

Furthermore,

$$G^{\mathrm{A}}_{s|s'}(z) = \int_0^\infty e^{-izt'}G^{\mathrm{A}}_{s|s'}(0|t')dt' = \left[\int_0^\infty e^{iz^*t'}G^{\mathrm{R}}_{s'|s}(t'|0)dt'\right]^* = G^{\mathrm{R}}_{s'|s}(z^*)^* \tag{142}$$

is analytic in the lower half plane of z. In contrast,

$$G^{\mathrm{K}}_{s|s'}(z) = \int_{-\infty}^{\infty} e^{-izt'} G^{\mathrm{K}}_{s|s'}(0|t')dt' = -\left[\int_{-\infty}^{\infty} e^{iz^*t'} G^{\mathrm{K}}_{s'|s}(t'|0)dt'\right]^* = -G^{\mathrm{K}}_{s'|s}(z^*)^* \tag{143}$$

has in general nonanalytic features in both half planes.
For the Keldysh components of the free Green function we find

$$g^{\mathrm{R}}(\omega) = -i\int_{-\infty}^{\infty} e^{i\omega t}\Theta(t)e^{-i\epsilon t}dt = \frac{1}{\omega - \epsilon + i\eta}, \tag{144}$$

where the positive infinitesimal η regularizes the integral. Similarly,

$$g^{\mathrm{A}}(\omega) = \frac{1}{\omega - \epsilon - i\eta}. \tag{145}$$

In order to obtain a Fourier integral of g^{K} that converges at least in the sense of distributions, we need to assure time translational invariance of $g^{\mathrm{K}}(t,t')$. Therefore we use the additional assumption

$$[\rho_0, H_0] = 0, \tag{146}$$

so that the occupation matrix n commutes with the matrix u_0 of free single-particle time evolution. In this case,

$$g^{\mathrm{K}}(t,t') = -iu_0(t,-\infty)(1+\zeta 2n)u_0(-\infty,t') = -i(1+\zeta 2n)u_0(t,t'), \tag{147}$$

which yields

$$g^{\mathrm{K}}(\omega) = -i2\pi\delta(\omega-\epsilon)(1+\zeta 2n) \tag{148}$$

$$= (1+\zeta 2n)\left[g^{\mathrm{R}}(\omega) - g^{\mathrm{A}}(\omega)\right]. \tag{149}$$

Due to time translational invariance, the Fourier transformed self-energy is

$$\Sigma(\omega'|\omega) = 2\pi\delta(\omega'-\omega)\Sigma(\omega), \tag{150}$$

with $\Sigma(\omega) = \Sigma(t'=0|\omega) = \Sigma(\omega|t=0)$. Due to the δ-functions, the frequency convolutions in Dyson's equation reduce to simple multiplications,

$$G(\omega) = g(\omega) + g(\omega)\Sigma(\omega)G(\omega), \tag{151}$$

with solution

$$G^{\mathrm{R}}(\omega) = \frac{1}{\omega - \epsilon - \Sigma^{\mathrm{R}}(\omega)} = G^{\mathrm{A}}(\omega)^{\dagger} \tag{152}$$

$$G^{\mathrm{K}}(\omega) = G^{\mathrm{R}}(\omega)\Sigma^{\mathrm{K}}(\omega)G^{\mathrm{A}}(\omega). \tag{153}$$

Here we used that

$$(G^R\Sigma^{\mathrm{K}}_{\mathrm{ini}}G^{\mathrm{A}})(t|t') \xrightarrow{t_0\to-\infty} 0 \tag{154}$$

due to the decay of long-term correlations. Since $\Sigma^{\mathrm{R}}(t',t) \sim \Theta(t'-t)$ it follows as for the Green function that $\Sigma^{\mathrm{R}}(z)$ is analytic in the upper half plane of the complex frequency z.
Finally, the Fourier transformed interaction vertex reads

$$\bar{\nu}_{\xi_1'\xi_2'\xi_1\xi_2} = \begin{cases} 2\pi\delta(\omega_1'+\omega_2'-\omega_1-\omega_2)\frac{\bar{v}_{s_1's_2's_1s_2}}{2}, & l_1'+l_2'+l_1+l_2 \text{ odd}, \\ 0, & l_1'+l_2'+l_1+l_2 \text{ even}. \end{cases} \tag{155}$$

2.8 Keldysh formalism in equilibrium

Suppose the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$ does not depend on time. Let the initial density matrix describe grand canonical equilibrium with temperature $T = \beta^{-1}$ and chemical potential μ. As $\hat{H}$ conserves particle number,

$$[\hat{H}, \hat{N}] = 0, \tag{156}$$

the density matrix is stationary,

$$\hat{\rho}(t) = \hat{\rho}(t_0) = \hat{\rho}_{\text{eq}} = \frac{1}{\mathcal{Z}} e^{-\beta(\hat{H}-\mu\hat{N})}, \qquad \mathcal{Z} = \operatorname{Tr} e^{-\beta(\hat{H}-\mu\hat{N})}, \tag{157}$$

where

$$\hat{N} = \sum_s \hat{a}_s^\dagger \hat{a}_s \tag{158}$$

is the operator of total particle number.

Fluctuation dissipation theorem. We exploit that the density matrix has the structure of a time evolution in imaginary time direction,

$$\hat{a}_s(t)_{\text{H}} \hat{\rho}_{\text{eq}} = \frac{1}{\mathcal{Z}} e^{i\hat{H}(t-t_0)} \hat{a}_s e^{-i\hat{H}(t-t_0)} e^{-\beta(\hat{H}-\mu\hat{N})} \tag{159}$$

$$= \frac{1}{\mathcal{Z}} e^{-\beta(\hat{H}-\mu\hat{N})} e^{i\hat{H}(t-i\beta-t_0)} e^{-\beta\mu\hat{N}} \hat{a}_s e^{\beta\mu\hat{N}} e^{-i\hat{H}(t-i\beta-t_0)} \tag{160}$$

$$= e^{\beta\mu} \hat{\rho}_{\text{eq}} \hat{a}_s(t - i\beta)_{\text{H}}. \tag{161}$$

For the single-particle Green function follows [9]

$$G^<_{ss'}(t-t') = -i\zeta \operatorname{Tr} \hat{a}_{s'}^\dagger(t')_{\text{H}} \hat{a}_s(t)_{\text{H}} \hat{\rho}_{\text{eq}} \tag{162}$$

$$= -i\zeta e^{\beta\mu} \operatorname{Tr} \hat{a}_s(t - i\beta)_{\text{H}} \hat{a}_{s'}^\dagger(t')_{\text{H}} \hat{\rho}_{\text{eq}} \tag{163}$$

$$= \zeta e^{\beta\mu} G^>_{ss'}(t - i\beta - t'). \tag{164}$$

We change to frequency representation,

$$G^<(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} G^<(t) dt \tag{165}$$

$$= \zeta e^{-\beta(\omega-\mu)} \int_{-\infty-i\beta}^{\infty-i\beta} e^{i\omega t} G^>(t) dt \tag{166}$$

$$= \zeta e^{-\beta(\omega-\mu)} G^>(\omega). \tag{167}$$

In the last step we used that $G^>(t)$ is analytic in t and vanishes for $t \to \infty$. By use of

$$G^{\text{R}}(\omega) - G^{\text{A}}(\omega) = G^>(\omega) - G^<(\omega) \tag{168}$$

$$G^{\text{K}}(\omega) = G^>(\omega) + G^<(\omega), \tag{169}$$

this equation characterizing thermal equilibrium can be transformed to the so called fluctuation dissipation theorem

$$G^{\text{K}}(\omega) = [1 + \zeta 2 n_\zeta(\omega)] \left[G^{\text{R}}(\omega) - G^{\text{A}}(\omega) \right], \tag{170}$$

where

$$n_\zeta(\omega) = \frac{1}{e^{\beta(\omega-\mu)} - \zeta} \tag{171}$$

denotes the Bose or Fermi function. As $G^{\mathrm{A}} = G^{\mathrm{R}\,\dagger}$, all single-particle Green functions can be derived from G^{R} alone in thermal equilibrium.
An analogous equation for the self-energy follows from

$$\begin{aligned} G^{\mathrm{R}}(\omega)\Sigma^{\mathrm{K}}(\omega)G^{\mathrm{A}}(\omega) &= G^{\mathrm{K}}(\omega) && (172)\\ &= \left[1 + \zeta 2n_\zeta(\omega)\right] G^{\mathrm{R}}(\omega) \left[G^{\mathrm{A}\,-1}(\omega) - G^{\mathrm{R}\,-1}(\omega)\right] G^{\mathrm{A}}(\omega) && (173)\\ &= \left[1 + \zeta 2n_\zeta(\omega)\right] G^{\mathrm{R}}(\omega) \left[\Sigma^{\mathrm{R}}(\omega) - \Sigma^{\mathrm{A}}(\omega)\right] G^{\mathrm{A}}(\omega), && (174) \end{aligned}$$

namely

$$\Sigma^{\mathrm{K}}(\omega) = \left[1 + \zeta 2n_\zeta(\omega)\right] \left[\Sigma^{\mathrm{R}}(\omega) - \Sigma^{\mathrm{A}}(\omega)\right]. \tag{175}$$

It can as well be Keldysh back rotated to

$$\Sigma^{<}(\omega) = \zeta e^{-\beta(\omega-\mu)}\Sigma^{>}(\omega). \tag{176}$$

By similar arguments, one can also establish equilibrium conditions for multi-particle Green functions [10]. Their structure is however more complicated.
Diagrammatic expansion. As ρ_{eq} is not of the uncorrelated form described in Eqs. (45, 46), Wick's theorem cannot be directly applied. However, the density matrix can be written

$$\hat{\rho}_{\mathrm{eq}} = \frac{1}{\mathcal{Z}} e^{-i\hat{H}(t_0 - i\beta - t_0)} e^{\beta\mu\hat{N}} = \frac{1}{\mathcal{Z}} \hat{u}(t_0 - i\beta, t_0) e^{\beta\mu\hat{N}}. \tag{177}$$

Therefore,

$$\begin{aligned} \langle A\rangle(t) &= \mathrm{Tr}\, \hat{A}(t)_{\mathrm{H}} \rho_{\mathrm{eq}} && (178)\\ &= \frac{1}{\mathcal{Z}} \mathrm{Tr}\, \hat{u}(t_0 - i\beta, t_0)\hat{u}(t_0, t)A(t)\hat{u}(t, t_0)e^{\beta\mu\hat{N}} && (179)\\ &= \frac{1}{\mathcal{Z}} \mathrm{Tr}\, \hat{u}_0(t_0 - i\beta, t_0)\hat{u}_{\mathrm{I}}(t_0 - i\beta, t_0)\hat{u}_{\mathrm{I}}(t_0, t)A(t)_0\hat{u}_{\mathrm{I}}(t, t_0)e^{\beta\mu\hat{N}} && (180)\\ &= \frac{\mathcal{Z}_0}{\mathcal{Z}} \mathrm{Tr}\, \hat{u}_{\mathrm{I}}(t_0 - i\beta, t_0)\hat{u}_{\mathrm{I}}(t_0, t)A(t)_0\hat{u}_{\mathrm{I}}(t, t_0)\hat{\rho}_{\mathrm{eq},0}, && (181) \end{aligned}$$

where

$$\begin{aligned} \rho_{\mathrm{eq},0} &= \frac{1}{\mathcal{Z}_0} e^{-\beta(\hat{H}_0 - \mu\hat{N})} = \frac{1}{\mathcal{Z}_0} \hat{u}_0(t_0 - i\beta, t_0) e^{\beta\mu\hat{N}}, && (182)\\ \mathcal{Z}_0 &= \mathrm{Tr}\, e^{-\beta(\hat{H}_0 - \mu\hat{N})}. && (183) \end{aligned}$$

$\hat{\rho}_{\mathrm{eq},0}$ is of the form required for the application of Wick's theorem. This motivates us to add a third branch to the time contour which leads from t_0 to $t_0 - i\beta$ and which we label by $j = \mathrm{im}$ since it goes in (negative) imaginary time direction. We agree that $\hat{\mathcal{T}}_{\mathrm{c}}$ places operators with index $j = \mathrm{im}$ left of operators with $j = \mathrm{b}$ and $j = \mathrm{f}$. It arranges the im-operators such that the imaginary part of time becomes more and more negative from right to left. Then we can write

$$\langle A\rangle(t) = \frac{\mathcal{Z}_0}{\mathcal{Z}} \mathrm{Tr}\, \hat{\mathcal{T}}_{\mathrm{c}}\, e^{-i\left(\int_{t_0}^{\infty}[V^{\mathrm{f}}(\tau)_0 - V^{\mathrm{b}}(\tau)_0]d\tau + \int_{t_0}^{t_0 - i\beta} V^{\mathrm{im}}(\tau)_0 d\tau\right)} A^{\mathrm{f}}(t)_0 \rho_{\mathrm{eq},0}, \tag{184}$$

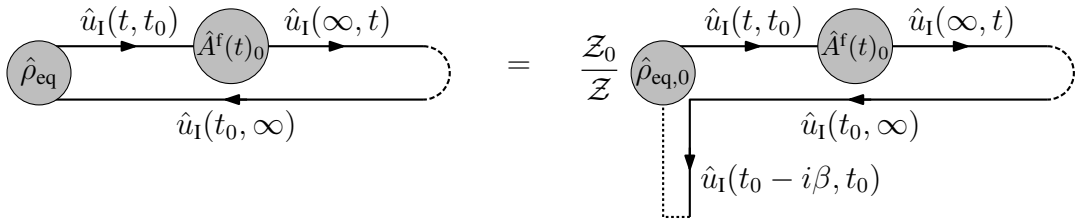


Fig. 9: *Pictorial representation of Eq. (184). We chose the vertical direction to signify time evolution in imaginary direction. The dotted line connecting the end of that time evolution with* $\hat{\rho}_{\rm eq,0}$ *just visualizes the cycle of the trace operation.*

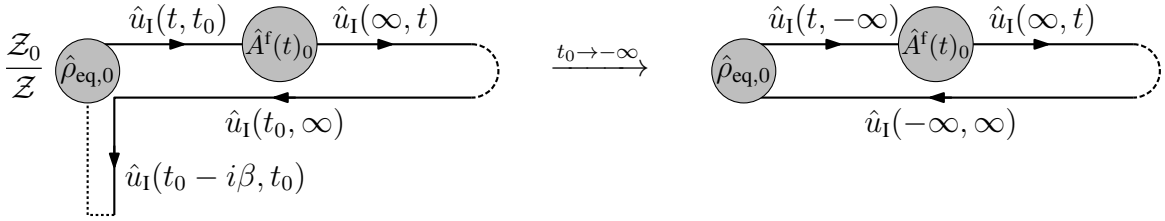


Fig. 10: *If inelastic processes lead to the decay of long-term correlations, the imaginary-time branch of the contour can be omitted (together with the factor* $\mathcal{Z}_0/\mathcal{Z}$*) in the limit* $t_0\to-\infty$*. This means in total, that* $\hat{\rho}_{\rm eq}$ *has been simply replaced by* $\hat{\rho}_{\rm eq,0}$*.*

as depicted in Fig. 9. Similarly, the n-particle Green function is

$$\begin{aligned}&G(x_1,\ldots,x_n|x_1',\ldots,x_n')\\&=(-i)^n\frac{\mathcal{Z}_0}{\mathcal{Z}}\,\mathrm{Tr}\,\hat{\mathcal{T}}_{\rm c}\,e^{-i\left(\int_{t_0}^{\infty}[V^{\rm f}(\tau)_0-V^{\rm b}(\tau)_0]d\tau+\int_{t_0}^{t_0-i\beta}V^{\rm im}(\tau)_0d\tau\right)}\hat{a}_{x_1}\ldots\hat{a}_{x_n}\hat{a}^\dagger_{x_n'}\ldots\hat{a}^\dagger_{x_1'}\hat{\rho}_{\rm eq,0}.\end{aligned}\tag{185}$$

Wick's theorem holds in identical form in this situation. When we expand the exponential in powers of the interaction, we can again represent all contributions by diagrams. The diagram rules are similar but take care of the third branch of the contour. As before, diagrams which are not connected to the external fields need not be considered. Together with the prefactor $\mathcal{Z}_0/\mathcal{Z}$ they produce a global factor

$$\begin{aligned}&\frac{\mathcal{Z}_0}{\mathcal{Z}}\,\mathrm{Tr}\,\hat{\mathcal{T}}_{\rm c}\,e^{-i\left(\int_{t_0}^{\infty}[V^{\rm f}(\tau)_0-V^{\rm b}(\tau)_0]d\tau+\int_{t_0}^{t_0-i\beta}V^{\rm im}(\tau)_0d\tau\right)}\hat{\rho}_{\rm eq,0}\\&=\frac{\mathcal{Z}_0}{\mathcal{Z}}\,\mathrm{Tr}\,\hat{u}_{\rm I}(t_0-i\beta,t_0)\hat{u}_{\rm I}(t_0,\infty)\hat{u}_{\rm I}(\infty,t_0)\hat{\rho}_{\rm eq,0}g &(186)\\&=\frac{1}{\mathcal{Z}}\,\mathrm{Tr}\,\hat{u}_0(t_0-i\beta,t_0)\hat{u}_{\rm I}(t_0-i\beta,t_0)e^{\beta\mu\hat{N}} &(187)\\&=\frac{1}{\mathcal{Z}}\,\mathrm{Tr}\,\hat{u}(t_0-i\beta,t_0)e^{\beta\mu\hat{N}} &(188)\\&=1. &(189)\end{aligned}$$

Often inelastic processes lead to the decay of long-term correlations. Then diagrammatic contributions connecting the external operators $\hat{a}_x,\hat{a}^\dagger_{x'}$ of a Green function with the imaginary time contour $\hat{u}_{\rm I}(t_0-i\beta,t_0)$ will vanish in the limit $t_0\to-\infty$. Consequently, the Green function can be computed from the remaining diagrams that do not involve the imaginary time contour,

$$G(x_1,\ldots,x_n|x_1',\ldots,x_n')=(-i)^n\,\mathrm{Tr}\,\hat{\mathcal{T}}_{\rm c}\,e^{-i\int_{t_0}^{\infty}[V^{\rm f}(\tau)_0-V^{\rm b}(\tau)_0]d\tau}\hat{a}_{x_1}\ldots\hat{a}_{x_n}\hat{a}^\dagger_{x_n'}\ldots\hat{a}^\dagger_{x_1'}\hat{\rho}_{\rm eq,0}.\tag{190}$$

Thus we can simply use the standard two-branch contour in this case and replace ρ_{eq} with $\rho_{\text{eq},0}$, compare Fig. 10.

Relation to Matsubara Green functions. Matsubara formalism [6] is tailored to study the properties of a system in thermal equilibrium. Its basic idea is to position all operators in question on the imaginary part of the time contour and to obtain results for real times by analytic continuation. The forward and backward branch of the contour are then no longer needed. The single-particle Matsubara Green function is defined for the (relative) time $\tau = -\operatorname{Im} t$ in the interval from 0 to β. Fourier transformation leads to the function $G^{\text{Matsu}}_{ss'}(\omega_n)$ with discrete Matsubara frequencies

$$\omega_n = \begin{cases} 2n\pi T, & \text{for Bosons,} \\ (2n+1)\pi T, & \text{for Fermions,} \end{cases} \qquad n \text{ integer.} \tag{191}$$

$G^{\text{Matsu}}(\omega_n)$ is related to $G^{\text{R}}(\omega)$ and $G^{\text{A}}(\omega)$ by analytic continuation. This can be seen from the Lehmann representation , which is achieved by evaluating the Green function in a multi-particle eigenbasis of $\hat{H}$ and $\hat{\rho}_{\text{eq}}$. Let $|n\rangle$ denote an orthonormal basis of common multi-particle eigenstates of $\hat{H}$ and $\hat{N}$,

$$\hat{H}\,|n\rangle = E_n\,|n\rangle\,, \qquad \hat{N}\,|n\rangle = N_n\,|n\rangle\,. \tag{192}$$

We use it to study the frequency dependent retarded Green function $G^{\text{R}}(z)$ with frequency z from the complex upper half plane,

$$G^{\text{R}}_{ss'}(z) = -i\int_0^\infty dt\, e^{izt}\,\operatorname{Tr}\frac{e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathcal{Z}}\left[\hat{a}_s(t)_{\text{H}}, \hat{a}^\dagger_{s'}(0)_{\text{H}}\right]_{-\zeta} \tag{193}$$

$$= -i\sum_{n,m}\frac{e^{-\beta(E_n-\mu N_n)}}{\mathcal{Z}}\int_0^\infty dt\, e^{izt}\Big[\langle n|a_s(t)_{\text{H}}|m\rangle\langle m|a^\dagger_{s'}(0)_{\text{H}}|n\rangle - \zeta\langle n|a^\dagger_{s'}(0)_{\text{H}}|m\rangle\langle m|a_s(t)_{\text{H}}|n\rangle\Big] \tag{194}$$

$$= -i\sum_{n,m}\frac{e^{-\beta(E_n-\mu N_n)}}{\mathcal{Z}}\int_0^\infty dt\,\left[e^{i(E_n-E_m+z)t}a_{s,nm}a^*_{s',nm} - \zeta e^{i(E_m-E_n+z)t}a_{s,mn}a^*_{s',mn}\right] \tag{195}$$

$$= \sum_{n,m}\frac{e^{-\beta(E_n-\mu N_n)}}{\mathcal{Z}}\left[\frac{a_{s,nm}a^*_{s',nm}}{E_n-E_m+z} - \zeta\frac{a_{s,mn}a^*_{s',mn}}{E_m-E_n+z}\right], \tag{196}$$

where

$$a_{s,nm} = \langle n|\hat{a}_s|m\rangle = \langle m|\hat{a}^\dagger_s|n\rangle^*. \tag{197}$$

Note that the integral converges only for z with positive imaginary part. $G^{\text{R}}(\omega)$ with real frequency ω is obtained as limit from above by replacing z with $\omega + i\eta$. The Lehmann representation of the advanced Green function is formally identical, but allows only for z from the lower half plane. The Lehmann representation of the Matsubara Green function turns out to be again formally identical but with imaginary frequency $i\omega_n$. We thus have

$$G^{\text{Matsu}}(\omega_n) = \begin{cases} G^{\text{R}}(z = i\omega_n + \mu), & \text{if } \omega_n > 0, \\ G^{\text{A}}(z = i\omega_n + \mu), & \text{if } \omega_n < 0. \end{cases} \tag{198}$$

The shift of the frequency by μ results from the fact that Matsubara formalism defines "time evolution" to be generated by $\hat{H} - \mu\hat{N}$ instead of just $\hat{H}$. This trick is technically convenient and without influence on expectation values and correlation functions of number conserving operators.

2.9 Generating functionals for Green and vertex functions

The n-particle Green functions can be identified as expansion coefficients of a so called generating functional. This makes it possible to derive features of the Green functions from the properties of the generating functional. For instance, the flow equations of the functional renormalization group (see lecture B7 by V. Meden) can be derived from generating functionals. A rather detailed discussion of generating functionals in Matsubara formalism can be found in [6]. Generating functionals within Keldysh formalism are studied in [11, 12].

Generating functional for Green functions. In order to motivate the form of the generating functional, we consider adding a source term $\hat{H}_\theta(t)$ to the Hamiltonian,

$$\hat{H}'(t) = \hat{H}(t) + \hat{H}_\theta(t) \tag{199}$$

$$\hat{H}_\theta(t) = \sum_s \left[\bar{\theta}_s(t)\hat{a}_s + \hat{a}_s^\dagger\theta_s(t)\right]. \tag{200}$$

Here, $\theta_s(t)$ and $\bar{\theta}_s(t)$ are conjugated numbers. $\theta_s(t)$ can be chosen complex. For Fermions however it is convenient to choose $\theta_s(t)$ and $\bar{\theta}_s(t)$ as conjugated generators of a Grassmann algebra. These generators anti-commute with each other and with the annihilators and creators. $\hat{H}_\theta(t)$ operates on a generalized Fock space then which allows for Grassmann coefficients. Now

$$1 = \mathrm{Tr}\,\hat{u}(t_0,\infty)\hat{u}(\infty,t_0)\hat{\rho}_0 \tag{201}$$

$$= \mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,\exp\left[-i\int_{t_0}^{\infty}\sum_j \mathrm{sgn}(j)\left(\hat{H}^j(\tau) + \sum_s\left[\bar{\theta}_s(\tau)\hat{a}_s^j(\tau) + \hat{a}_s^{\dagger\,j}(\tau)\theta_s(\tau)\right]\right)d\tau\right]\hat{\rho}_0. \tag{202}$$

In the last expression, the time dependence of $\hat{a}_s^j(\tau)$ is only formal and indicates the position of the operator under contour ordering. Let us now formally allow $\theta_s(\tau)$ to take different values on the forward and on the backward branch. We define the functional

$$\mathcal{G}[\theta,\bar{\theta}] = \mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,\exp\left[-i\int_{t_0}^{\infty}\sum_j \left(\mathrm{sgn}(j)\hat{H}^j(\tau) + \sum_s\left[\bar{\theta}_s^j(\tau)\hat{a}_s^j(\tau) + \hat{a}_s^{\dagger\,j}(\tau)\theta_s^j(\tau)\right]\right)d\tau\right]\hat{\rho}_0. \tag{203}$$

As seen above, $\mathcal{G}$ satisfies the normalization condition

$$\mathcal{G}[\theta,\bar{\theta}]\big|_{\substack{\theta^\mathrm{b}=-\theta^\mathrm{f}\\ \bar{\theta}^\mathrm{b}=-\bar{\theta}^\mathrm{f}}} = 1. \tag{204}$$

The functional can be used to generate Green functions. For instance

$$\left.\frac{\delta^2\mathcal{G}[\theta,\bar{\theta}]}{\delta\bar{\theta}_r^i(t)\delta\theta_{r'}^{i'}(t')}\right|_{\substack{\theta=0\\ \bar{\theta}=0}} = \mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,e^{-i\int_{t_0}^{\infty}\sum_j \mathrm{sgn}(j)\hat{H}^j(\tau)d\tau}[-i\hat{a}_r^i(t)][-i\zeta\hat{a}_{r'}^{\dagger\,i'}(t')]\hat{\rho}_0 \tag{205}$$

$$= -\zeta\,\mathrm{Tr}\,\hat{\mathcal{T}}_\mathrm{c}\,\hat{a}_r^i(t)_\mathrm{H}\hat{a}_{r'}^{\dagger\,i'}(t')_\mathrm{H}\rho_0 \tag{206}$$

$$= -i\zeta G_{r|r'}^{i|i'}(t|t') \tag{207}$$

(except at $t = t'$ where the derivative does not reproduce the rule to sort $a^\dagger$ to the left of a). The derivative with respect to $\theta^{i'}_{r'}(t')$ produced a factor $\zeta = -1$ for Fermions due to the anti-commutative behaviour of the Grassmann generators. More generally, the n-particle Green function is given by

$$G(x_1, \ldots, x_n | x'_1, \ldots, x'_n) = (\zeta i)^n \left. \frac{\delta^{2n} \mathcal{G}[\theta, \bar{\theta}]}{\delta\bar{\theta}_{x_1} \ldots \delta\bar{\theta}_{x_n} \delta\theta_{x'_n} \ldots \delta\theta_{x'_1}} \right|_{\substack{\theta=0 \\ \bar{\theta}=0}}. \tag{208}$$

The expansion of the generating functional in the fields $\theta, \bar{\theta}$ then reads

$$\mathcal{G}[\theta, \bar{\theta}] = 1 + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!^2} G(x_1, \ldots, x_n | x'_1, \ldots, x'_n) \bar{\theta}_{x_1} \ldots \bar{\theta}_{x_n} \theta_{x'_n} \ldots \theta_{x'_1}, \tag{209}$$

with summation convention for repeated multi-indices.
As an example how features of the Green functions can be derived from the generating functional we study the normalization condition. After Keldysh rotation from $\theta^{\mathrm{f,b}}$ to $\theta^{1,2}$ via

$$\theta_z = K_{zx} \theta_x \tag{210}$$

the normalization condition reads

$$\mathcal{G}[\theta, \bar{\theta}]\big|_{\substack{\theta^2=0 \\ \bar{\theta}^2=0}} = 1. \tag{211}$$

This entails

$$G^{1\ldots1|1\ldots1} = (\zeta i)^n \left. \frac{\delta^{2n} \mathcal{G}}{\delta\bar{\theta}^1 \ldots \delta\bar{\theta}^1 \delta\theta^1 \ldots \delta\theta^1} \right|_{\substack{\theta=0 \\ \bar{\theta}=0}} = (\zeta i)^n \left. \frac{\delta^{2n} \left(\mathcal{G}|_{\theta^2=0=\bar{\theta}^2}\right)}{\delta\bar{\theta}^1 \ldots \delta\bar{\theta}^1 \delta\theta^1 \ldots \delta\theta^1} \right|_{\substack{\theta^1=0 \\ \bar{\theta}^1=0}} = 0, \tag{212}$$

which is a generalization of our result for the patch Green function in the single particle case.
Further generating functionals. Other classes of functions can be generated from appropriate functionals as well. For instance, the connected n-particle Green function $G_{\mathrm{c}}(x_1, \ldots, x_n | x'_1, \ldots, x'_n)$ is defined as sum of all connected n-particle diagrams. A diagram is called connected, if any two (internal or external) vertices are connected by a sequence of lines and other vertices. It can be shown [6], that the connected Green functions are generated by $\ln \mathcal{G}$,

$$G_{\mathrm{c}}(x_1, \ldots, x_n | x'_1, \ldots, x'_n) = (\zeta i)^n \left. \frac{\delta^{2n} \mathcal{W}[\theta, \bar{\theta}]}{\delta\bar{\theta}_{x_1} \ldots \delta\bar{\theta}_{x_n} \delta\theta_{x'_n} \ldots \delta\theta_{x'_1}} \right|_{\substack{\theta=0 \\ \bar{\theta}=0}}, \tag{213}$$

$$\mathcal{W}[\theta, \bar{\theta}] = \ln G[\theta, \bar{\theta}]. \tag{214}$$

The functional renormalization group is often formulated for the so called one-particle irreducible vertex functions $\gamma(x'_1, \ldots, x'_n | x_1, \ldots, x_n)$. These are the multi-particle generalizations of the self energy. $\gamma(x'_1, \ldots, x'_n | x_1, \ldots, x_n)$ is defined to be the sum of all one-particle irreducible n-particle diagrams. It is convenient to define the value of such a diagram as

$$D(x'_1, \ldots, x'_n | x_1, \ldots, x_n) = i^{1-n} \frac{\zeta^P \zeta^{n_{\mathrm{loop}}}}{2^{n_{\mathrm{eq}}} S} \left(\prod i\bar{\nu} \right) \left(\prod g \right) \tag{215}$$

with the conventional factor i^{1-n}. This factor assures that the one-particle irreducible diagram consisting only of a bare n-particle vertex (if the Hamiltonian contains an n-particle interaction) is equal in value to that vertex.

It can be shown, that a Legendre transformation of $\mathcal{W}$ leads to the generating functional of the one-particle irreducible vertex functions. Define

$$\psi_x[\theta,\bar{\theta}] = \frac{\delta\left(i\mathcal{W}[\theta,\bar{\theta}]\right)}{\delta\bar{\theta}_x}, \tag{216}$$

$$\bar{\psi}_{x'}[\theta,\bar{\theta}] = \zeta\frac{\delta\left(i\mathcal{W}[\theta,\bar{\theta}]\right)}{\delta\theta_{x'}}, \tag{217}$$

and suppose this can be solved for $\theta[\psi,\bar{\psi}], \bar{\theta}[\psi,\bar{\psi}]$. Then the so called effective action

$$\Gamma[\psi,\bar{\psi}] = \left[\bar{\theta}_x\psi_x + \bar{\psi}_{x'}\theta_{x'} - i\mathcal{W}[\theta,\bar{\theta}] - \bar{\psi}_{x'}g^{-1}(x',x)\psi_x\right]_{\substack{\theta[\psi,\bar{\psi}]\\ \bar{\theta}[\psi,\bar{\psi}]}} \tag{218}$$

generates the vertex functions via

$$\gamma(x'_1,\ldots,x'_n|x_1,\ldots,x_n) = -\zeta^n\frac{\delta^{2n}\Gamma[\psi,\bar{\psi}]}{\delta\bar{\psi}_{x'_1}\ldots\delta\bar{\psi}_{x'_n}\delta\psi_{x_n}\ldots\delta\psi_{x_1}}\bigg|_{\substack{\psi=0\\ \bar{\psi}=0}}. \tag{219}$$

Coherent state functional integral representation. For explicit calculations it is often convenient to work with generating functionals in the coherent state functional integral representation. We sketch briefly, how this representation is achieved for the functional $\mathcal{G}$. The time span from t_0 to $t_{\max}$ (with $t_{\max}\to\infty$) is divided into N steps, $N\to\infty$,

$$t_k = t_0 + k\Delta, \tag{220}$$

$$\Delta = \frac{t_{\max}-t_0}{N}, \tag{221}$$

such that

$$\mathcal{G}[\theta,\bar{\theta}] = \lim_{t_{\max}\to\infty}\lim_{N\to\infty}\operatorname{Tr}\hat{u}^{\mathrm{b}}_\theta(t_0,t_1)\ldots\hat{u}^{\mathrm{b}}_\theta(t_{N-1},t_N)1\hat{u}^{\mathrm{f}}_\theta(t_N,t_{N-1})\ldots\hat{u}^{\mathrm{f}}_\theta(t_1,t_0)\hat{\rho}_0, \tag{222}$$

with

$$\hat{u}^{\mathrm{f}}_\theta(t_k,t_{k-1}) = \hat{\mathcal{T}}_{\mathrm{c}}\exp\left[-i\int_{t_{k-1}}^{t_k}\sum_j\left(\hat{H}(\tau)+\sum_s\left[\bar{\theta}^{\mathrm{f}}_s(\tau)\hat{a}_s(\tau)+\hat{a}^\dagger_s(\tau)\theta^{\mathrm{f}}_s(\tau)\right]\right)d\tau\right] \tag{223}$$

$$\approx e^{-i\Delta\left(\hat{H}(t_k)+\sum_s\left[\bar{\theta}^{\mathrm{f}}_s(t_k)\hat{a}_s+\hat{a}^\dagger_s\theta^{\mathrm{f}}_s(t_k)\right]\right)} \tag{224}$$

$$\hat{u}^{\mathrm{b}}_\theta(t_{k-1},t_k) \approx e^{i\Delta\left(\hat{H}(t_k)-\sum_s\left[\bar{\theta}^{\mathrm{b}}_s(t_k)\hat{a}_s+\hat{a}^\dagger_s\theta^{\mathrm{b}}_s(t_k)\right]\right)}. \tag{225}$$

Now one introduces coherent state resolutions of unity between any two operators in that expression. The coherent states $|\phi\rangle$ [with $\phi=(\phi_{s_1},\phi_{s_2},\ldots)$] are given by [6]

$$|\phi\rangle = e^{\zeta\sum_s\phi_s\hat{a}^\dagger_s}|\mathrm{vac}\rangle \tag{226}$$

and satisfy

$$\hat{a}_s|\phi\rangle = \phi_s|\phi\rangle, \tag{227}$$

$$\langle\phi|\hat{a}^\dagger_s = \bar{\phi}_s\langle\phi| \tag{228}$$

for all s. For Bosons, the ϕ_s are complex numbers. For Fermions, they are generators of a Grassmann algebra. The coherent state resolution of unity reads

$$\hat{1} = \int \left(\prod_s \frac{d\bar{\phi}_s d\phi_s}{\mathcal{N}} e^{-\bar{\phi}_s \phi_s} \right) |\phi\rangle\langle\phi| \,, \tag{229}$$

and the trace of a particle number conserving operator is

$$\mathrm{Tr}\, \hat{A} = \int \left(\prod_s \frac{d\bar{\phi}_s d\phi_s}{\mathcal{N}} e^{-\bar{\phi}_s \phi_s} \right) \langle\phi|\hat{A}|\zeta\phi\rangle, \tag{230}$$

where

$$\mathcal{N} = \begin{cases} 2\pi i, & \text{for Bosons,} \\ 1, & \text{for Fermions.} \end{cases} \tag{231}$$

The resulting expression for the generating functional is finally

$$\mathcal{G}[\theta, \bar{\theta}] = \int \mathcal{D}\phi \mathcal{D}\bar{\phi}\, e^{i\left(A[\phi,\bar{\phi}] - \bar{\theta}_x \phi_x - \bar{\phi}_{x'} \theta_{x'}\right)}. \tag{232}$$

Here, we used the abbreviation

$$\int \mathcal{D}\phi \mathcal{D}\bar{\phi} \ldots = \lim_{t_{\max}\to\infty} \lim_{N\to\infty} \int \prod_s \prod_{j=\mathrm{f,b}} \prod_{k=0}^{N} \frac{d\bar{\phi}^j_{k,s} d\phi^j_{k,s}}{\mathcal{N}} \ldots \tag{233}$$

The action

$$A[\phi, \bar{\phi}] = A_0[\phi, \bar{\phi}] + A_{\text{int}}[\phi, \bar{\phi}] \tag{234}$$

is given by

$$A_0[\phi, \bar{\phi}] = \bar{\phi}_{x'} g^{-1}(x', x) \phi_x \tag{235}$$

and

$$A_{\text{int}}[\phi, \bar{\phi}] = -\frac{1}{4} \bar{\nu}_{x'_1 x'_2 x_1 x_2} \bar{\phi}_{x'_1} \bar{\phi}_{x'_2} \phi_{x_2} \phi_{x_1}, \tag{236}$$

where

$$\bar{\nu}_{x'_1 x'_2 x_1 x_2} = \delta_{j'_1 j_1} \delta_{j'_2 j_1} \delta_{j_2 j_1} \,\mathrm{sgn}(j_1) \delta(t'_1 - t_1) \delta(t'_2 - t_1) \delta(t_2 - t_1) \bar{v}_{s'_1 s'_2 s_1 s_2}(t_1). \tag{237}$$

The coherent state representation of the effective action is

$$\Gamma[\psi, \bar{\psi}] = -i \ln \left[\frac{1}{\mathcal{Z}} \int \mathcal{D}\phi \mathcal{D}\bar{\phi}\, e^{i\left(A[\phi,\bar{\phi}] - A_0[\psi,\bar{\psi}] - \bar{\theta}_x \phi_x - \bar{\phi}_{x'} \theta_{x'} + \bar{\theta}_x \psi_x + \bar{\psi}_{x'} \theta_{x'}\right)} \right]_{\substack{\theta[\psi,\bar{\psi}] \\ \bar{\theta}[\psi,\bar{\psi}]}}. \tag{238}$$

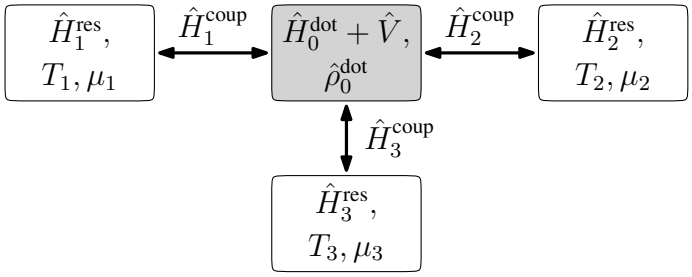


Fig. 11: *The system under consideration consists of an interacting quantum dot that is tunnel coupled to M non-interacting leads, here $M = 3$. At time t_0 the leads are prepared in thermal equilibrium, each with individual temperature and chemical potential.*

3 Nonequilibrium transport through quantum systems

3.1 Model for a quantum system coupled to leads

We consider a confined mesoscopic region which is coupled via barriers to several leads. This geometry is realized for example by a quantum dot or wire with quantum point contacts, embedded in a semi-conductor hetero structure; or by a carbon nanotube lying across metallic electrodes; or by a molecule coupled by ligands into a metallic break junction. Due to the smallness of the confined region, transport is phase coherent so that a quantum mechanical description is required. Furthermore, interactions have decisive influence on transport properties.
Let us for example focus on effects of the particle-particle interaction in a system of Fermionic particles. The confined region (called dot) is coupled to M noninteracting reservoirs, $M \geq 1$, which are supposed to be metallic with long-living quasi-particles and described by a single-particle Hamiltonian. The total Hamiltonian under consideration is then given by

$$\hat{H}(t) = \hat{H}_0^{\text{dot}}(t) + \hat{V} + \sum_{r=1}^{M} \left[\hat{H}_r^{\text{res}} + \hat{H}_r^{\text{coup}}(t) \right], \tag{239}$$

with

$$\hat{H}_0^{\text{dot}}(t) = \sum_{q',q} \epsilon_{q'q}(t) \hat{c}_{q'}^\dagger \hat{c}_q \tag{240}$$

$$\hat{V} = \frac{1}{4} \sum_{\substack{q_1,q_1' \\ q_2,q_2'}} \overline{v}_{q_1' q_2' q_1 q_2} \hat{c}_{q_1'}^\dagger \hat{c}_{q_2'}^\dagger \hat{c}_{q_2} \hat{c}_{q_1}, \tag{241}$$

$$\hat{H}_r^{\text{res}} = \sum_{k_r} \epsilon_{k_r} \hat{c}_{k_r}^\dagger \hat{c}_{k_r}, \tag{242}$$

$$\hat{H}_r^{\text{coup}}(t) = \sum_{k_r,q} J_{k_r q}(t) \hat{c}_{k_r}^\dagger \hat{c}_q + \text{H. c.} \tag{243}$$

in standard notation of second quantization. Here, q denotes single-particle states which are localized on the dot, while k_r denotes the single-particle eigenstates of the Hamiltonian of reservoir r. For homogeneous electron reservoirs, k is a multi-index comprising momentum

and spin. The single-particle matrix elements of the Hamiltonian are

$$\epsilon_{q'q}(t) = \langle q'|\hat{H}_0^{\text{dot}}(t)|q\rangle \tag{244}$$

$$\epsilon_{k_r} = \langle k_r|\hat{H}_r^{\text{res}}|k_r\rangle, \tag{245}$$

$$J_{k_r q}(t) = \langle k_r|\hat{H}_r^{\text{coup}}(t)|q\rangle = J_{qk_r}(t)^*, \tag{246}$$

while

$$\overline{v}_{q_1'q_2'q_1q_2} = \langle q_1'q_2'|\hat{V}|q_1q_2\rangle - \langle q_1'q_2'|\hat{V}|q_2q_1\rangle \tag{247}$$

denotes the antisymmetrized matrix elements of the two-particle interaction.
We assume that at some initial time t_0 the whole configuration is described by the product density matrix

$$\hat{\rho}(t_0) = \hat{\rho}_0 = \hat{\rho}_0^{\text{dot}} \otimes \hat{\rho}_{1,0}^{\text{res}} \otimes \ldots \otimes \hat{\rho}_{M,0}^{\text{res}}, \tag{248}$$

where

$$\hat{\rho}_{r,0}^{\text{res}} = \frac{e^{-(\hat{H}_r^{\text{res}} - \mu_r \hat{N}_r)/T_r}}{\operatorname{Tr} e^{-(\hat{H}_r^{\text{res}} - \mu_r \hat{N}_r)/T_r}}, \qquad r = 1, \ldots, M, \tag{249}$$

characterizes grand-canonical equilibrium of reservoir r with temperature T_r and chemical potential μ_r, and

$$\hat{N}_r = \sum_{k_r} \hat{c}_{k_r}^\dagger \hat{c}_{k_r} \tag{250}$$

is the operator of particle number in reservoir r. We denote the corresponding Fermi functions by

$$f_r(\omega) = \frac{1}{e^{(\omega-\mu_r)/T_r} + 1}. \tag{251}$$

A product density matrix arises naturally if dot and reservoirs are decoupled for $t < t_0$. Different temperatures or chemical potentials of the different reservoirs prepare a global nonequilibrium situation. The so prepared system is depicted schematically in Fig. 11.
For times $t > t_0$ the Hamiltonian $\hat{H}$ induces a unitary time evolution. If the reservoirs are infinite in size and are prepared with different temperatures or chemical potentials, the system will typically evolve towards a state of stationary transport. This state, as well as the time dependent transient behaviour before, can be studied within Keldysh formalism.

3.2 Reservoir self-energy

Consider a diagrammatic expansion of the multi-particle Green functions, where $\hat{H}_0^{\text{dot}} + \sum_r \hat{H}_r^{\text{res}}$ are taken as noninteracting part of the Hamiltonian, while $\hat{V} + \sum_r \hat{H}_r^{\text{coup}}$ serve as interaction. The two contributions to the interaction give rise to two types of vertices: two-particle vertices $\bar{\nu}$ representing the two-particle interaction $\hat{V}$ and one-particle vertices

$$\theta_{qk_r}^{j'j}(t', t) = \delta(t' - t)\delta_{j'j}\,\mathrm{sgn}(j) J_{qk_r}(t) \tag{252}$$

$$\theta_{k_r q}^{j'j}(t', t) = \delta(t' - t)\delta_{j'j}\,\mathrm{sgn}(j) J_{k_r q}(t) \tag{253}$$

$$= \theta_{qk_r}^{jj'}(t, t')^*, \tag{254}$$

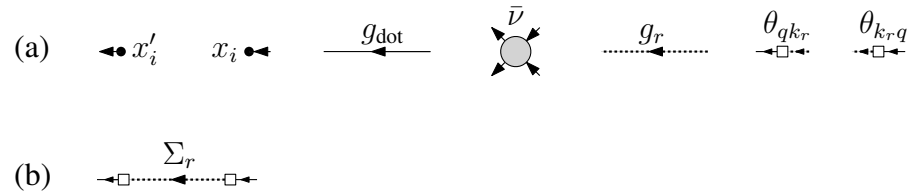


Fig. 12: *(a) The constituents of diagrams if the Green function is expanded in powers of $\hat{V}$ and of $\hat{H}_r^{\text{coup}}$. (b) Diagrammatic representation of the self-energy Σ_r describing exscursions to reservoir r.*

see Fig. 12 (a). The value of a diagram is given by

$$\frac{(-1)^P(-1)^{n_{\text{loop}}}}{2^{n_{\text{eq}}}S}\left(\prod i\bar{\nu}\right)\left(\prod\theta\right)\left(\prod g_{\text{dot}}\right)\left(\prod g_r\right). \tag{255}$$

Often one is interested in expectation values of operators that act merely on the dot and not on the reservoirs. They can be derived from Green functions with only dot states as external indices. The propagation between two dot states can contain an arbitrary number of temporal excursions into each reservoir. As the reservoirs are noninteracting, each such excursion is described by $\theta g_r\theta$, which is tunneling into the reservoir, free propagation in the reservoir, and tunneling back into the dot, compare Fig. 12 (b). Summing up all possible numbers of excursions to a given reservoir r yields a series which corresponds exactly to Dyson's equation with the self-energy $\Sigma_r = \theta g_r\theta$ having components

$$\Sigma_r{}^{j'|j}_{q'|q}(t'|t) = \int_{t_0}^{\infty} dt_1 dt'_1 \sum_{k_r}\sum_{i,i'=\text{f,b}} \theta^{j'i}_{q'k_r}(t',t_1) g^{i|i'}_{k_r}(t_1|t'_1)\theta^{i'j}_{k_rq}(t'_1,t) \tag{256}$$

$$= \operatorname{sgn}(j')\operatorname{sgn}(j)\sum_{k_r} J_{q'k_r}(t') g^{j'|j}_{k_r}(t'|t) J_{k_rq}(t), \tag{257}$$

$$\Sigma_r{}^{\text{R,K}}_{q'|q}(t'|t) = \sum_{k_r} J_{q'k_r}(t') g^{\text{R,K}}_{k_r}(t'|t) J_{k_rq}(t), \tag{258}$$

where the last line follows after Keldysh rotation. The advanced component can be obtained as adjoint of the retarded. Since the reservoir is in noninteracting thermal equilibrium, we have

$$g^{\text{R}}_{k_r}(t|t') = -i\Theta(t-t')e^{-i\epsilon_{k_r}(t-t')} \tag{259}$$

$$g^{\text{K}}_{k_r}(t|t') = -i\left[1-2f_r(\epsilon_{k_r})\right]e^{-i\epsilon_{k_r}(t-t')}. \tag{260}$$

The free Green function depends on k_r via ϵ_{k_r} only. Therefore we can write

$$\Sigma_r{}^{\text{R,K}}_{q'|q}(t'|t) = \frac{1}{2\pi}\int d\epsilon\, \Gamma^{(r)}_{q'q}(\epsilon;t',t) g^{\text{R,K}}_r(\epsilon;t'|t), \tag{261}$$

with the hybridization function

$$\Gamma^{(r)}_{q'q}(\epsilon;t',t) = 2\pi\sum_{k_r} J_{q'k_r}(t')\delta(\epsilon-\epsilon_{k_r})J_{k_rq}(t). \tag{262}$$

Note, that different conventions for the definition of $\Gamma^{(r)}$ exist in the literature. Often, the factor 2 on the right hand side of Eq. (262) is omitted. Additional confusion can arise in situations with several identical reservoirs, where both of the different definitions $\Gamma = \sum_r \Gamma^{(r)}$ and

$\Gamma = \Gamma^{(r)}$ are frequently used. In a stationary situation, $\Gamma^{(r)}_{q'q}(\epsilon; t', t) = \Gamma^{(r)}_{q'q}(\epsilon)$ does not depend on time, and we get

$$\Sigma^{\mathrm{R}}_{r q'|q}(\omega) = \frac{1}{2\pi} \int d\epsilon \, \frac{\Gamma^{(r)}_{q'q}(\epsilon)}{\omega - \epsilon + i\eta} \tag{263}$$

$$\Sigma^{\mathrm{K}}_{r q'|q}(\omega) = [1 - 2f_r(\omega)] \left[\Sigma^{\mathrm{R}}_{r q'|q}(\omega) - \Sigma^{\mathrm{A}}_{r q'|q}(\omega)\right] \tag{264}$$

$$= -i \, [1 - 2f_r(\omega)] \, \Gamma^{(r)}_{q'q}(\omega). \tag{265}$$

In the following, we discuss tow prominent examples for the reservoir self-energy.

Wide-band limit. In the wide-band limit one assumes that $\Gamma^{(r)}_{q'q}(\epsilon; t', t) = \Gamma^{(r)}_{q'q}(t', t)$ does not depend on ϵ. This implies in particular a continuous density of states in the reservoir and therefore requires reservoirs of infinite size. If we assume that $J_{k_r q}$ does not depend strongly on k_r, then the ϵ-dependence of $\Gamma^{(r)}$ is mainly due to the density of states. Consequently, the wide-band limit implies a constant density of states. In a one-dimensional system, for instance, this would be given in case of a linear dispersion

$$\epsilon_{k_r} = v_F k_r \tag{266}$$

with Fermi velocity v_F. In the wide-band limit, one finds

$$\Sigma^{\mathrm{R}}_r(t'|t) = \frac{1}{2\pi} \Gamma^{(r)}(t', t) \int d\epsilon \, \left[-i\Theta(t' - t) e^{-i\epsilon(t'-t)}\right] \tag{267}$$

$$= -i\delta(t' - t) \frac{\Gamma^{(r)}(t, t)}{2} \tag{268}$$

and

$$\Sigma^{\mathrm{K}}_r(t'|t) = \frac{1}{2\pi} \Gamma^{(r)}(t', t) \int d\epsilon \, \left(-i \, [1 - 2f_r(\epsilon)] \, e^{-i\epsilon(t'-t)}\right). \tag{269}$$

Now we use the identity

$$1 - 2f(\epsilon) = -2T \sum_{\omega_m} \frac{1}{i\omega_m - \epsilon + \mu}, \tag{270}$$

where the sum runs over the Matsubara frequencies $\omega_m = (2m + 1)\pi T$ with integers m. The sum is to be evaluated as principal value for $|\omega_m| \to \infty$. This leads to

$$\Sigma^{\mathrm{K}}_r(t'|t) = -\frac{i}{\pi} \Gamma^{(r)}(t', t) T_r \sum_{\omega^{(r)}_m} \int d\epsilon \, \frac{e^{-i\epsilon(t'-t)}}{\epsilon - \mu_r - i\omega^{(r)}_m} \tag{271}$$

which can be solved by residue theorem to yield

$$\Sigma^{\mathrm{K}}_r(t'|t) = -\Gamma^{(r)}(t', t) T_r \frac{e^{-i\mu_r(t'-t)}}{\sinh\left[(t' - t)\pi T_r\right]}. \tag{272}$$

This has to be evaluated as principal value when integrated over t' and t through $t' = t$. For the particular case $T_r = 0$ an expansion of the $\sinh$ leads to

$$\Sigma^{\mathrm{K}}_r(t'|t) = -\frac{1}{\pi} \Gamma^{(r)}(t', t) \frac{e^{-i\mu_r(t'-t)}}{t' - t}, \qquad \text{for} \quad T_r = 0. \tag{273}$$

In a stationary situation, $\Gamma^{(r)}(t',t) = \Gamma^{(r)}$ does not depend on time, and the wide band limit yields

$$\Sigma_r^{\mathrm{R}}(\omega) = -\frac{i}{2}\Gamma^{(r)} \tag{274}$$

$$\Sigma_r^{\mathrm{K}}(\omega) = -i\left[1 - 2f_r(\omega)\right]\Gamma^{(r)}. \tag{275}$$

Tight-binding chain. As a second example for the reservoir self-energy we consider a reservoir consisting of a semi-infinite tight binding chain which is coupled to the dot only with its first site. For simplicity we choose spinless particles, assume a stationary state, and work in frequency representation. If we suppress the index r indicating the reservoir number for convenience, the reservoir Hamiltonian is given by

$$\hat{H}_{\mathrm{res}} = -w\sum_{m=1}^{\infty}\hat{c}_m^\dagger\hat{c}_{m+1} + \mathrm{H.\,c.}, \tag{276}$$

where we have chosen a real hopping amplitude $(-w) < 0$. The operator $\hat{c}_m^\dagger$ generates a particle in the single-particle state $|x_m\rangle$ on the m-th site centered around

$$x_m = ma, \qquad m = 1, 2, \ldots, \tag{277}$$

with the lattice constant a. The states $|x_m\rangle$ shall constitute a complete orthonormal set of states. The single-particle eigenstates of the reservoir Hamiltonian are standing waves. Suitably normalized they form a complete orthonormal set,

$$|k\rangle = \sqrt{\frac{2a}{\pi}}\sum_{m=1}^{\infty}\sin(kx_m)\,|x_m\rangle\,, \qquad k \in \left(0, \frac{\pi}{a}\right), \tag{278}$$

$$\int_0^{\pi/a} dk\ |k\rangle\langle k| = \sum_{m=1}^{\infty}|x_m\rangle\langle x_m|\,, \tag{279}$$

$$\langle k'|k\rangle = \delta(k-k'), \tag{280}$$

with eigenenergies

$$\epsilon_k = -2w\cos(ka). \tag{281}$$

Since we couple only the first reservoir site to the dot,

$$J_{kq} = \langle k|\hat{H}^{\mathrm{coup}}|q\rangle = \langle k|x_1\rangle\,\langle x_1|\hat{H}^{\mathrm{coup}}|q\rangle = \sqrt{\frac{2a}{\pi}}\sin(ka)J_{x_1q}. \tag{282}$$

The retarded reservoir self-energy then reads

$$\Sigma^{\mathrm{R}}_{\mathrm{res}\,q'|q}(\omega) = \int dk\, J_{q'k}g_k^{\mathrm{R}}(\omega)J_{kq} \tag{283}$$

$$= J_{q'x_1}J_{x_1q}\frac{2}{\pi}\int_0^{\pi} d\kappa\,\frac{\sin^2(\kappa)}{\omega + 2w\cos(\kappa) + i\eta} \tag{284}$$

with $\kappa = ka$. First we solve the integral for real ω with $|\omega| > 2w$. Then we deduce the solution for $|\omega| < 2w$ by analytic continuation via the upper complex half plane of ω. For real ω with $|\omega| > 2w$ we use $\sin^2\kappa = (1-\cos 2\kappa)/2$ and

$$\int_0^\pi d\kappa\, \frac{\cos(m\kappa)}{1+\alpha\cos\kappa} = \frac{\pi}{\sqrt{1-\alpha^2}} \left(\frac{\sqrt{1-\alpha^2}-1}{\alpha}\right)^{|m|}, \quad \text{for } m \text{ integer and } -1<\alpha<1, \tag{285}$$

to find

$$\Sigma^{\mathrm{R}}_{\mathrm{res}\,q'|q}(\omega) = \frac{J_{q'x_1}J_{x_1q}}{w}\frac{\omega}{2w}\left[1-\sqrt{1-\left(\tfrac{2w}{\omega}\right)^2}\right], \qquad \text{for } \omega \text{ real with } |\omega|>2w. \tag{286}$$

Analytic continuation to the complex upper half plane yields

$$\Sigma^{\mathrm{R}}_{\mathrm{res}\,q'|q}(z) = \frac{J_{q'x_1}J_{x_1q}}{w}\left[\frac{z}{2w} - i\sqrt{1-\left(\tfrac{z}{2w}\right)^2}\right] \quad \text{for } \mathrm{Im}\, z>0, \tag{287}$$

where the branch-cut of $\sqrt{z}$ is defined to be on the negative real axis. Continuation from the upper half plane to the real axis results finally in

$$\Sigma^{\mathrm{R}}_{\mathrm{res}\,q'|q}(\omega) = \frac{J_{q'x_1}J_{x_1q}}{w} \times \begin{cases} \frac{\omega}{2w}\left[1-\sqrt{1-\left(\frac{2w}{\omega}\right)^2}\right], & \omega \text{ real, } |\omega|>2w, \\ \left[\frac{\omega}{2w} - i\sqrt{1-\left(\frac{\omega}{2w}\right)^2}\right], & \omega \text{ real, } |\omega|<2w. \end{cases} \tag{288}$$

From this result one can derive the hybridization function

$$\Gamma_{q'q}(\epsilon) = i\left[\Sigma^{\mathrm{R}}_{\mathrm{res}\,q'q}(\epsilon) - \Sigma^{\mathrm{R}}_{\mathrm{res}\,qq'}(\epsilon)^*\right] \tag{289}$$

$$= \Theta\big(2w-|\epsilon|\big)\, 2\frac{J_{q'x_1}J_{x_1q}}{w}\sqrt{1-\left(\tfrac{\epsilon}{2w}\right)^2}. \tag{290}$$

The wide-band limit of a hybridization that does not depend on energy is obtained in the limit

$$w\to\infty \quad \text{with} \quad \frac{J_{q'x_1}J_{x_1q}}{w} = \text{const.}, \tag{291}$$

in which

$$\Gamma_{q'q}(\epsilon) \to \Gamma_{q'q} \equiv 2\frac{J_{q'x_1}J_{x_1q}}{w}. \tag{292}$$

3.3 Dot occupancy, electric current, heat current

In this section we discuss, how typical observables can be computed from the Green functions.
Occupancy of dot states. The time dependent expectation value of the occupancy of the single-particle dot state q can be found from

$$\hat{c}_q^\dagger\hat{c}_q = \frac{1}{2}\left(\hat{c}_q\hat{c}_q^\dagger + \hat{c}_q^\dagger\hat{c}_q - \hat{c}_q\hat{c}_q^\dagger + \hat{c}_q^\dagger\hat{c}_q\right) = \frac{1}{2}\left(1-\left[\hat{c}_q,\hat{c}_q^\dagger\right]\right) \tag{293}$$

to be

$$\langle\hat{c}_q^\dagger\hat{c}_q\rangle(t) = \mathrm{Tr}(\hat{c}_q^\dagger\hat{c}_q)(t)_{\mathrm{H}}\hat{\rho}_0 = \frac{1}{2}\left[1-iG^{\mathrm{K}}_{q|q}(t|t)\right]. \tag{294}$$

Electric current. If each particle (electron) carries the charge e, the expected electric current I_r leaving reservoir r is

$$I_r(t) = -e\,\mathrm{Tr}\,\frac{d}{dt}\hat{N}_r(t)_{\mathrm{H}}\hat{\rho}_0 = -ie\,\mathrm{Tr}[\hat{H},\hat{N}_r](t)_{\mathrm{H}}\hat{\rho}_0, \tag{295}$$

where

$$[\hat{H}(t),\hat{N}_r] = [\hat{H}_r^{\mathrm{coup}}(t),\hat{N}_r] = -\sum_{k_r,q}\left[J_{k_rq}(t)\hat{c}^\dagger_{k_r}\hat{c}_q - \mathrm{H.c.}\right]. \tag{296}$$

From

$$\mathrm{Tr}\,c^\dagger_{k_r}(t)_{\mathrm{H}}\hat{c}_q(t)_{\mathrm{H}}\hat{\rho}_0 = \frac{1}{2}\,\mathrm{Tr}\left[c^\dagger_{k_r}(t)_{\mathrm{H}},\hat{c}_q(t)_{\mathrm{H}}\right]\hat{\rho}_0 = -\frac{i}{2}G^{\mathrm{K}}_{qk_r}(t|t) \tag{297}$$

follows

$$I_r(t) = e\,\mathrm{Re}\sum_{k_r,q}G^{\mathrm{K}}_{qk_r}(t|t)J_{k_rq}(t). \tag{298}$$

Now we can use

$$G^{\mathrm{K}}_{qk_r}(t|t) = \sum_{q'}\int_{t_0}^{t}dt'\Big[G^{\mathrm{R}}_{qq'}(t|t')J_{q'k_r}(t')g^{\mathrm{K}}_{k_r}(t'|t) + G^{\mathrm{K}}_{qq'}(t|t')J_{q'k_r}(t')g^{\mathrm{A}}_{k_r}(t'|t)\Big], \tag{299}$$

which follows from the fact that the reservoirs are noninteracting. It can be derived by the same diagrammatics as applied at the beginning of section 3.2 in order to motivate the concept of a reservoir self-energy. We find

$$I_r(t) = e\,\mathrm{Re}\int_{t_0}^{t}dt'\;\mathrm{Tr}\left[G^{\mathrm{R}}(t|t')\Sigma^{\mathrm{K}}_r(t'|t) + G^{\mathrm{K}}(t|t')\Sigma^{\mathrm{A}}_r(t'|t)\right], \tag{300}$$

where trace and multiplication abbreviate summation over dot quantum numbers. The current leaving the dot follows from conservation of total particle number,

$$I_{\mathrm{dot}}(t) = -\sum_r I_r(t) = -e\,\mathrm{Re}\int_{t_0}^{t}dt'\;\mathrm{Tr}\left[G^{\mathrm{R}}(t|t')\Sigma^{\mathrm{K}}_{\mathrm{res}}(t'|t) + G^{\mathrm{K}}(t|t')\Sigma^{\mathrm{A}}_{\mathrm{res}}(t'|t)\right], \tag{301}$$

with $\Sigma_{\mathrm{res}} = \sum_r \Sigma_r$ denoting the total reservoir self-energy.
In the stationary case, $I_r(t) \equiv I_r$ can be expressed in frequency representation as

$$I_r = \frac{e}{4\pi}\int d\omega\;\mathrm{Tr}\left\{\left[G^{\mathrm{R}}(\omega) - G^{\mathrm{A}}(\omega)\right]\Sigma^{\mathrm{K}}_r(\omega) - G^{\mathrm{K}}(\omega)\left[\Sigma^{\mathrm{R}}_r(\omega) - \Sigma^{\mathrm{A}}_r(\omega)\right]\right\} \tag{302}$$

$$= \frac{-ie}{4\pi}\int d\omega\;\mathrm{Tr}\,\Gamma_r(\omega)\left\{[1-2f_r(\omega)]\left[G^{\mathrm{R}}(\omega) - G^{\mathrm{A}}(\omega)\right] - G^{\mathrm{K}}(\omega)\right\}. \tag{303}$$

We note in passing that in case of global equilibrium, $f_r \equiv f$, the fluctuation dissipation theorem

$$G^{\mathrm{K}}(\omega) = [1-2f(\omega)]\left[G^{\mathrm{R}}(\omega) - G^{\mathrm{A}}(\omega)\right] \tag{304}$$

directly leads to $I_r = 0$, as expected. By use of

$$G^{\mathrm{K}} = G^{>} + G^{<}, \qquad G^{\mathrm{R}} - G^{\mathrm{A}} = G^{>} - G^{<}, \tag{305}$$

we can as well express the current in contour index representation,

$$I_r = \frac{ie}{4\pi} \int d\omega \ \mathrm{Tr}\, \Gamma_r(\omega) \left\{ f_r(\omega) G^{>}(\omega) + [1 - f_r(\omega)]\, G^{<}(\omega) \right\}. \tag{306}$$

Here, the term $\mathrm{Tr}\, f(\omega)\Gamma(\omega)\,[i\mathcal{G}^{>}(\omega)]$ can be interpreted as transitions of particles from the reservoir into the dot. The mean occupancy of a singe-particle state in the reservoir is given by $f(\omega)$, the transition rate to the dot is $\Gamma(\omega)$ and availability of free states in the dot is described by the quasi-hole weight $[i\mathcal{G}^{>}(\omega)]$. In the same spirit, the term $\mathrm{Tr}\,[-i\mathcal{G}^{<}(\omega)]\,\Gamma(\omega)[1 - f(\omega)]$ describes transitions of particles from the dot into the reservoir.
In the particular case of a noninteracting stationary system, the only self-energy contribution stems from the reservoirs. Then

$$G^{\mathrm{K}}(\omega) = G^{\mathrm{R}}(\omega) \sum_r \Sigma_r^{\mathrm{K}}(\omega) G^{\mathrm{A}}(\omega) \tag{307}$$

$$= -i \sum_r [1 - 2f_r(\omega)]\, G^{\mathrm{R}}(\omega)\Gamma_r(\omega) G^{\mathrm{A}}(\omega), \tag{308}$$

$$G^{\mathrm{R}}(\omega) - G^{\mathrm{A}}(\omega) = G^{\mathrm{R}}(\omega) \left[G^{\mathrm{A}\,-1}(\omega) - G^{\mathrm{R}\,-1}(\omega) \right] G^{\mathrm{A}}(\omega) \tag{309}$$

$$= -iG^{\mathrm{R}}(\omega) \sum_r \Gamma_r(\omega) G^{\mathrm{A}}(\omega), \tag{310}$$

results in a formula of the Landauer type [13]

$$I_r = \frac{e}{2\pi} \int d\omega \sum_{r'} [f_r(\omega) - f_{r'}(\omega)]\, \mathrm{Tr}\, \Gamma_r(\omega) G^{\mathrm{R}}(\omega) \Gamma_{r'}(\omega) G^{\mathrm{A}}(\omega) \tag{311}$$

$$= \frac{e}{2\pi} \int d\omega \sum_{r'} [f_r(\omega) - f_{r'}(\omega)]\, T_{rr'}(\omega), \tag{312}$$

with

$$T_{rr'}(\omega) = \mathrm{Tr}\, \Gamma_r(\omega) G^{\mathrm{R}}(\omega) \Gamma_{r'}(\omega) G^{\mathrm{A}}(\omega) \tag{313}$$

serving as transmission coefficient for scattering from reservoir r to reservoir r'.
There is a special interacting case in which a current formula of the Landauer type can be derived as well. Suppose that we study a stationary transport situation with only two reservoirs, a left and a right one, $r = \mathrm{L}, \mathrm{R}$. Let us assume that the hybridization matrices $\Gamma_r(\omega)$ are multiples of each other. This means they can be written

$$\Gamma_r(\omega) = c_r \Gamma(\omega), \tag{314}$$

with real valued coefficients c_r satisfying $c_{\mathrm{L}} + c_{\mathrm{R}} = 1$. Such a situation arises for instance in the single-impurity Anderson model, see lecture B3 by T. A. Costi. Exploiting current conservation, $I_{\mathrm{L}} = -I_{\mathrm{R}}$, the general current formula (303) can now be simplified to [14]

$$I_{\mathrm{L}} = c_{\mathrm{R}} I_{\mathrm{L}} - c_{\mathrm{L}} I_{\mathrm{R}} \tag{315}$$

$$= \frac{-ie}{4\pi} \int d\omega\, c_{\mathrm{L}} c_{\mathrm{R}}\, \mathrm{Tr}\, \Gamma(\omega) \left[2f_{\mathrm{R}}(\omega) - 2f_{\mathrm{L}}(\omega) \right] \left[G^{\mathrm{R}}(\omega) - G^{\mathrm{A}}(\omega) \right] \tag{316}$$

$$= \frac{e}{2\pi} \int d\omega \ [f_{\mathrm{L}}(\omega) - f_{\mathrm{R}}(\omega)]\, T_{\mathrm{LR}}(\omega), \tag{317}$$

with transmission coefficient

$$T_{\mathrm{LR}'}(\omega) = 2\pi c_{\mathrm{L}} c_{\mathrm{R}} \,\mathrm{Tr}\,\Gamma(\omega)\rho(\omega), \tag{318}$$

where

$$\rho(\omega) = \frac{i}{2\pi}\left[G^{\mathrm{R}}(\omega) - G^{\mathrm{A}}(\omega)\right] \tag{319}$$

denotes the spectral function.

Heat current. According to the first law of thermodynamics, the heat δQ_r leaving reservoir r can be computed from the change in energy $dE_r = d\,\mathrm{Tr}\,\hat{H}_r^{\mathrm{res}}(t)_{\mathrm{H}}\rho_0$ of that reservoir and from the work $\mu_r d\,\mathrm{Tr}\,\hat{N}_r(t)_{\mathrm{H}}\rho_0$ related to particles exchange. The heat current leaving the reservoir is hence

$$I_r^Q(t) = \frac{\delta Q_r}{dt} = -\,\mathrm{Tr}\,\frac{d}{dt}\left[\hat{H}_r^{\mathrm{res}}(t)_{\mathrm{H}} - \mu_r\hat{N}_r(t)_H\right]\rho_0 = -i\,\mathrm{Tr}\left[\hat{H}, \hat{H}_r^{\mathrm{res}} - \mu_r\hat{N}_r\right](t)_{\mathrm{H}}\rho_0. \tag{320}$$

Identical steps as for the particle current yield

$$I_r^Q(t) = \mathrm{Re}\sum_{k_r,q,q'}\int_{t_0}^{t} dt' J_{q'k_r}(t')J_{k_rq}(t)(\epsilon_{k_r} - \mu_r)\left[G^{\mathrm{R}}_{qq'}(t|t')g^{\mathrm{K}}_{k_r}(t'|t) + G^{\mathrm{K}}_{qq'}(t|t')g^{\mathrm{A}}_{k_r}(t'|t)\right] \tag{321}$$

$$= \mathrm{Re}\int\frac{d\epsilon}{2\pi}\int_{t_0}^{t} dt'(\epsilon - \mu_r)\,\mathrm{Tr}\,\Gamma^{(r)}(\epsilon;t',t)\left[G^{\mathrm{R}}(t|t')g_r^{\mathrm{K}}(\epsilon;t'|t) + G^{\mathrm{K}}(t|t')g_r^{\mathrm{A}}(\epsilon;t'|t)\right]. \tag{322}$$

In the stationary case the corresponding expression in frequency representation is

$$I_r^Q = \frac{-i}{4\pi}\int d\omega\,(\omega - \mu_r)\,\mathrm{Tr}\,\Gamma_r(\omega)\left\{[1 - 2f_r(\omega)]\left[G^{\mathrm{R}}(\omega) - G^{\mathrm{A}}(\omega)\right] - G^{\mathrm{K}}(\omega)\right\}. \tag{323}$$

In situations, where the electric current can be computed from a transmission coefficient, the analogue result for the heat current is [15]

$$I_r^Q = \frac{1}{2\pi}\int d\omega\,(\omega - \mu_r)\sum_{r'}\left[f_r(\omega) - f_{r'}(\omega)\right]T_{rr'}(\omega). \tag{324}$$

3.4 Transport coefficients

Let us consider a stationary transport situation with two reservoirs, $r = \mathrm{L},\mathrm{R}$ (left, right). Electric current and heat current are functions of $\mu_{\mathrm{L,R}}$ and $T_{\mathrm{L,R}}$. For given mean values $\mu = (\mu_{\mathrm{L}} + \mu_{\mathrm{R}})/2$ and $T = (T_{\mathrm{L}} + T_{\mathrm{R}})/2$ we consider the currents as functions of the bias voltage $V = (\mu_L - \mu_R)/e$ and the temperature difference $\tau = T_{\mathrm{L}} - T_{\mathrm{R}}$. The linear transport properties of the configuration can then be derived from the coefficients [15]

$$\begin{pmatrix} L_{11} & L_{12} \\ L_{21} & L_{22}\end{pmatrix} = \left.\begin{pmatrix} \frac{\partial I_{\mathrm{L}}(V,\tau)}{\partial V} & \frac{\partial I_{\mathrm{L}}(V,\tau)}{\partial \tau} \\ \frac{\partial I_{\mathrm{L}}^Q(V,\tau)}{\partial V} & \frac{\partial I_{\mathrm{L}}^Q(V,\tau)}{\partial \tau}\end{pmatrix}\right|_{\substack{V=0\\ \tau=0}}. \tag{325}$$

In particular, the linear conductance is

$$G = \left.\frac{\partial I_{\mathrm{L}}(V,\tau)}{\partial V}\right|_{\substack{V=0\\ \tau=0}} = L_{11}. \tag{326}$$

From

$$\left(\frac{\partial V}{\partial \tau}\right)_{I_\mathrm{L}} \left(\frac{\partial \tau}{\partial I_\mathrm{L}}\right)_{V} \left(\frac{\partial I_\mathrm{L}}{\partial V}\right)_{\tau} = -1 \tag{327}$$

and

$$(\tau = 0 \text{ and } I_\mathrm{L} = 0) \quad \Leftrightarrow \quad (\tau = 0 \text{ and } V = 0) \tag{328}$$

follows for the thermoelectric power (Seebeck coefficient)

$$S = -\left.\frac{\partial V(I_\mathrm{L}, \tau)}{\partial \tau}\right|_{\substack{I_\mathrm{L}=0 \\ \tau=0}} = L_{12}/L_{11}. \tag{329}$$

S indicates, which bias voltage counterbalances the driving effect of a temperature difference on the electric current. Furthermore, $I_\mathrm{L}^Q(I_\mathrm{L}, \tau) = I_\mathrm{L}^Q\big(V(I_\mathrm{L}, \tau), \tau\big)$ yields

$$\left(\frac{\partial I_\mathrm{L}^Q}{\partial \tau}\right)_{I_\mathrm{L}} = \left(\frac{\partial I_\mathrm{L}^Q}{\partial \tau}\right)_{V} + \left(\frac{\partial I_\mathrm{L}^Q}{\partial V}\right)_{\tau} \left(\frac{\partial V}{\partial \tau}\right)_{I_\mathrm{L}}. \tag{330}$$

For the electronic contribution to the thermal conductance this leads to

$$K = -\left.\frac{\partial I_\mathrm{L}^Q(I_\mathrm{L}, \tau)}{\partial \tau}\right|_{\substack{I_\mathrm{L}=0 \\ \tau=0}} = -L_{22} + L_{21}S = -L_{22} + L_{12}L_{21}/L_{11}. \tag{331}$$

We derive formulae for the coefficients L_{ij} for the cases in which

$$I_\mathrm{L} = \frac{e}{2\pi} \int d\omega \, [f_\mathrm{L}(\omega) - f_\mathrm{R}(\omega)] \, T_\mathrm{LR}(\omega). \tag{332}$$

We use

$$f_r(\omega) = \frac{1}{e^{(\omega-\mu_r)/T_r} + 1}, \qquad f(\omega) = \frac{1}{e^{(\omega-\mu)/T} + 1} \tag{333}$$

to find

$$\left.\frac{\partial}{\partial V} (f_\mathrm{L} - f_\mathrm{R})\right|_{\substack{V=0 \\ \tau=0}} = -2e\frac{\partial f}{\partial \omega}. \tag{334}$$

Together with

$$[f_\mathrm{L} - f_\mathrm{R}]_{\substack{V=0 \\ \tau=0}} = 0 \tag{335}$$

follows

$$L_{11} = -\frac{e^2}{\pi} \int d\omega \, \frac{\partial f(\omega)}{\partial \omega} T_\mathrm{LR}^\mathrm{eq}(\omega), \tag{336}$$

where $T_\mathrm{LR}^\mathrm{eq}(\omega) = [T_\mathrm{LR}(\omega)]_{\substack{V=0 \\ \tau=0}}$. The factor $(-\partial f/\partial \omega)$ in the integrand is a peak of weight one centered around $\omega = \mu$ and with width in the order of T. The linear conductance $G = L_{11}$ is hence determined by the weight of the transmission coefficient $T_\mathrm{LR}^\mathrm{eq}(\omega)$ in the corresponding region. Similarly,

$$\left.\frac{\partial}{\partial \tau} (f_\mathrm{L} - f_\mathrm{R})\right|_{\substack{V=0 \\ \tau=0}} = -2\frac{\omega - \mu}{T}\frac{\partial f}{\partial \omega} \tag{337}$$

yields

$$L_{12} = -\frac{e}{\pi T}\int d\omega\,(\omega-\mu)\frac{\partial f(\omega)}{\partial\omega}T_{\mathrm{LR}}^{\mathrm{eq}}(\omega). \tag{338}$$

Obviously, non-vanishing thermoelectric power $S = L_{12}/L_{11}$ requires an asymmetry of $T_{\mathrm{LR}}^{\mathrm{eq}}(\omega)$ around the chemical potential. As I_{L}^{Q} differs from I_{L} only by a factor $(\omega-\mu_{\mathrm{L}})/e$ in the integrand, the corresponding coefficients for the heat current are

$$L_{21} = TL_{12}, \tag{339}$$

$$L_{22} = -\frac{1}{\pi T}\int d\omega\,(\omega-\mu)^2\frac{\partial f(\omega)}{\partial\omega}T_{\mathrm{LR}}^{\mathrm{eq}}(\omega). \tag{340}$$

We have seen that all coefficients in questions can be computed by suitable frequency integrals of the equilibrium transmission coefficient.

References

[1] J. Rammer, *Quantum Field Theory of Nonequilibrium States* (Cambridge University Press, Cambridge 2011)

[2] H. J. W. Haug and A.-P. Jauho, *Quantum Kinetics in Transport and Optics of Semiconductors* (Springer, Berlin, 2008)

[3] P. Danielewicz, Ann. Phys. (N.Y.) **152**, 239 (1984)

[4] J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986)

[5] A. Kamenev and A. Levchenko, Adv. Phys. **58**, 197 (2009)

[6] J. W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, Reading, 1988)

[7] L. V. Keldysh, Soviet Physics JETP **20**, 1018 (1965) [Zh. Eksp. Teor. Fiz. **47**, 1515 (1964)]

[8] A. I. Larkin and Y. N. Ovchinnikov, Sov. Phys. JETP **41**, 960 (1975) [Zh. Eksp. Teor. Fiz. **68**, 1915 (1975)]

[9] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (Benjamin, Menlo Park, 1962)

[10] S. G. Jakobs, M. Pletyukhov, and H. Schoeller, J. Phys. A: Math. Theor. **43**, 103001 (2010)

[11] K.-C. Chou, Z.-B. Su, B.-L. Hao, and L. Yu, Phys. Rep. **118**, 1 (1985)

[12] R. Gezzi, T. Pruschke, and V. Meden, Phys. Rev. B **75**, 045324 (2007)

[13] R. Landauer, Philos. Mag. **21**, 863 (1970)

[14] Y. Meir and N. S. Wingreen, Phys. Rev. Lett. **68**, 2512 (1992)

[15] U. Sivan and Y. Imry, Phys. Rev. B **33**, 551 (1986)

B 3 Numerical Renormalization Group for Quantum Impurities[1]

T. A. Costi
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

This lecture deals with a particular implementation of the renormalization group (RG) idea: Wilson's non-perturbative numerical renormalization group (NRG) method for quantum impurity models[1]. This version of the NRG, which will be described in detail below, was originally developed in the context of the Kondo model of magnetic impurities (such as Fe or Mn) in non-magnetic metals (such as Cu, Au, Ag etc)[2]. The Kondo model is defined by the Hamiltonian

$$H_{\rm KM} = J\vec{S}.\vec{s}_0 + \sum_{k\sigma} \varepsilon_{k\sigma} c^+_{k\sigma} c_{k\sigma}, \tag{1}$$

and describes a localized impurity spin $\vec{S}$ interacting antiferromagnetically ($J > 0$) with the conduction electrons of the host, via their spin-density $\vec{s}_0$ at the impurity site, see Sec. 2 for details. Unlike the case of non-magnetic impurities, or potential scatterers (see Lecture A5 by Mavropoulos), magnetic impurities have internal dynamical degrees of freedom which result in inelastic scattering of conduction electrons. This makes the Kondo problem, the scattering of electrons from magnetic impurities, a genuine many-body correlation problem. Wilson used the NRG to solve the many-body Hamiltonian (1) and demonstrated conclusively that a $S = 1/2$ magnetic impurity embedded in a non-magnetic metal has its magnetic moment completely screened by the surrounding conduction electrons, provided the temperature is sufficiently low, namely for $T \ll T_K$, where $T_K = \sqrt{JN_{\rm F}} e^{-1/JN_{\rm F}}$ is a dynamically generated low energy scale called the Kondo scale (see Sec. 2). This pioneering work established the formalism and gave a detailed analysis of the fixed points and thermodynamics of the Kondo model, and, later, also of the Anderson impurity model (see Sec. 2). The NRG has since been applied to many more quantum impurity models[2, 3, 4]. In addition, it has been extended to the calculation of equilibrium dynamical and transport properties [5, 6, 7, 8, 9, 10], e.g., dynamical susceptibilities, resistivities and thermopower or the conductance through quantum dots [11, 12], thereby making the NRG a useful tool for interpreting experiments which probe these quantities.

Two challenges for the NRG at present, are, (i), to extend it to more realistic multi-orbital and multi-channel models (e.g., for use in realistic modeling of materials), and, (ii), to extend it to the transient and non-equilibrium steady state response of quantum impurity systems. Recent progress and ideas in these two directions is outlined in Sec. 6.

The outline of this lecture is as follows: quantum impurity models are introduced in Sec. 2 and the Anderson impurity and Kondo models are described. The closely related spin-boson model is also briefly described (see also functional and real-time RG techniques in Lectures B7 and B5 by Meden and Schoeller, respectively). Wilson's NRG method is described in Sec. 3, and the calculation of physical properties is outlined in Sec. 4. In Sec. 5 we describe the recently introduced complete basis set and its use in constructing the full density matrix. Applications to thermodynamics and Green functions are given. An outline of some recent developments using the NRG is given in Sec. 6, and, Sec. 7 summarizes with possible future directions.

[2]Most readers will be familiar with the application of the renormalization group to critical phenomena and obtaining accurate values for critical exponents, see Appendix C. The NRG for quantum impurity models focuses on obtaining the complete spectrum of eigenvalues and eigenstates of such models on all energy scales and hence is more general in its aims, as will become clear in this Lecture

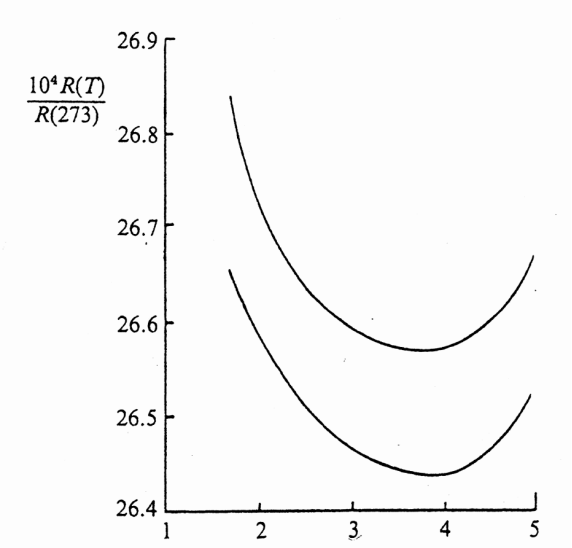


Fig. 1: *Resistivity $R(T)$ versus temperature $T[K]$ of two samples of "pure"* Au *showing the first observation of the resistivity minimum [17] . The expected behaviour of $R(T)$ for a pure metal with weak static disorder is a T^5 term due to phonons and a saturation to a constant value, ρ_0, at $T = 0$ due to static disorder. The former is seen in the experiment, but at low temperature an additional logarithmically increasing contribution is also found.*

2 Quantum impurity models

Quantum impurity models describe systems where the many-body interaction (usually a Coulomb or exchange interaction) acts at one or only a few sites, the "impurity", and the impurity is coupled to a large system, the bath, consisting of a macroscopically large number of non-interacting particles. These particles can be either bosons (e.g. phonons, magnons, photons, particle-hole pairs etc) or fermions (e.g. electrons in the conduction band, fermionic ^{40}K atoms in an optical lattice, etc). The "impurity" may be a real impurity, such as an Fe impurity (in Au), or a two-level atom (coupled to the electromagnetic field), or, just a confined region behaving like an artificial atom, as in the case of semiconductor quantum dots (coupled to leads). It may also simply represent the lowest two quantum mechanical states of a system with a double-well potential, as in the case of quantum tunneling between macroscopic fluxoid states in a superconducting quantum interference device, which can be used to realize a qubit for quantum computation. Two magnetic impurities in a non-magnetic metal at a distance R apart, interacting via the RKKY indirect exchange $J_{\rm RKKY}$ may also be regarded as a quantum impurity system [13]. Analogues of this in nanostructures, such as double quantum dots attached to leads also exist. The transfer of electrons between donor and acceptor molecules in photosynthesis and other biological processes may also be approximately described in terms of a two-state system coupled to environmental degrees of freedom (the solvent). Concrete models describing the above situations go under the names of (isotropic and anisotropic) single and multi-channel Kondo models, the Anderson impurity model and the dissipative two-state system [14, 15]. They describe a large number of physical systems of current experimental and theoretical interest. Quantum impurity models are also of relevance in the study of correlated lattice models, such as the Hubbard or Kondo lattice models, since the latter are often well approximated, via the dynamical mean field theory, by a local impurity model embedded in a medium which has to be determined self-consistently [16].

Historically, interest in quantum impurities arose when magnetic impurities were found to be present, albeit in very low concentrations, even in apparently very pure metals such as Au or Ag. In particular, measurements of the resistivity of Au as early as the 1930's showed an unexpected minimum at low temperature (Fig. 1). The puzzle of the resistivity minimum was resolved by

Kondo in 1964, who showed that a small concentration $c_{\rm imp}$ of *magnetic* impurities modelled by Eq. (1) gives rise to an additional temperature dependent term in the resistivity of the form $\rho_{\rm K} = -c_{\rm imp}\, b\, \ln(T/D)$, which increases with decreasing temperature. The balance between the decreasing phonon contribution behaving as $\rho_{\rm phonon} = aT^5$ and the increasing Kondo contribution gives rise to the observed resistivity minimum. The logarithmic contribution to the resistivity, found by Kondo in perturbation theory, cannot hold down to $T = 0$ as the total scattering remains finite in this limit (unitarity limit). Wilson's non-perturbative NRG provides a way to obtain the correct behaviour of the resistivity $\rho(T)$ from high temperatures through a crossover regime at $T \sim T_{\rm K}$ all the way down to zero temperature [see Fig. 10 (a) showing the analogous quantity for a Kondo correlated quantum dot, the conductance $G(T)$].
The general form of the Hamiltonian for any quantum impurity system is given by

$$H = H_{\rm imp} + H_{\rm int} + H_{\rm bath}, \tag{2}$$

where $H_{\rm imp}$ describes the impurity, a small quantum mechanical system with only a few degrees of freedom $H_{\rm bath}$ represents the bath, and $H_{\rm int}$ is the interaction between the two.
We next consider explicit examples which will appear in this and several other lectures.

Anderson and Kondo impurity models: linear chain form

The prototype model for strongly correlated systems is the single-band non-degenerate Anderson model [18, 19],

$$H_{AM} = \underbrace{\sum_{\sigma} \varepsilon_d n_{d\sigma} + U n_{d\uparrow} n_{d\downarrow}}_{H_{\rm imp}} + \underbrace{\sum_{k\sigma} V_{kd}(c^\dagger_{k\sigma} d_\sigma + d^\dagger_\sigma c_{k\sigma})}_{H_{\rm int}} + \underbrace{\sum_{k\sigma} \epsilon_k c^\dagger_{k\sigma} c_{k\sigma}}_{H_{\rm bath}}. \tag{3}$$

The first two terms describe the impurity, represented here by a non-degenerate s-level of energy ε_d (see Sec. 6 for generalizations). Electrons in the local level are subject to a Coulomb repulsion U which acts between spin-up and spin-down electrons. The local level hybridizes with the Bloch states of a non-interacting s-wave conduction band, the last term in H_{AM}, with amplitude $V_{\mathbf{k}d}$. The properties of the model are determined by the hybridization function

$$\Delta(\omega) = \pi \sum_k |V_{kd}|^2 \delta(\omega - \varepsilon_k), \tag{4}$$

which, like the conduction density of states $\rho(\omega) = \sum_k \delta(\omega - \varepsilon_k)$, will in general be a complicated function of energy. In cases where the interest is in the very low energy physics, it is a good approximation to set $\Delta(\omega) \approx \Delta(\varepsilon_F) \equiv \Delta$. In applications to pseudogap systems [20] or to effective quantum impurities in dynamical mean field theory, the full frequency dependence has to be retained. In applications to quantum dots, the impurity is attached to two baths, the left and right leads, as shown in Fig. 2. The Anderson model also provides an approximate description of the the low temperature behaviour of nanoscale size (ca. 100nm or smaller) quantum dots exhibiting the Kondo effect, see Fig. 2. [3]

[3] Although such dots are attached to two baths (the left and right leads), for a single level on the dot, only the even combination of left and right lead states couples to the dot. When several levels on the dot are active in transport, one will have a two-channel multi-orbital Anderson model with intra- and inter-orbital Coulomb interactions playing a role (e.g. Hund's exchange).

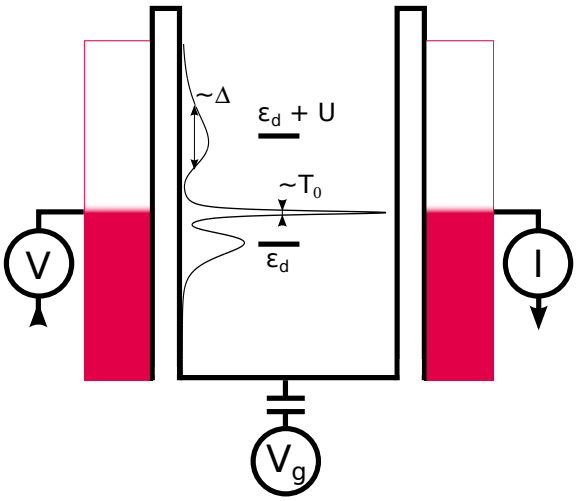


Fig. 2: *A quantum dot with charging energy $U \gg \Delta$ and level energy ε_d connected to left/right leads $H_{\alpha=L/R} = \sum_k \epsilon_{k\alpha\sigma} c^\dagger_{k\alpha\sigma} c_{k\alpha\sigma}$ via tunnel barriers. The gate voltage $V_g \sim \varepsilon_d$ allows changing ε_d relative to ϵ_F and thereby the dot occupation n_d from $n_d = 1$ for $\varepsilon_d = -U/2$ (Kondo regime) to $n_d = 0$ through a mixed valence regime with $n_d \approx 0.5$ for $\varepsilon_d \approx 0$. [21]*

Closely related to the Anderson model, is the Kondo model, which was briefly mentioned in the introduction. We write it's Hamiltonian as

$$H_{KM} = -g\mu_B B S_z + J\vec{S}.\vec{s}_0 + \sum_{k\sigma} \varepsilon_k c^+_{k\sigma} c_{k\sigma}, \tag{5}$$

where we included a magnetic field term $H_{\rm imp} = -g\mu_B B S_z$ to indicate the impurity spin $\vec{S}$ (taken here to be a $S = 1/2$ for simplicity), which interacts via an exchange interaction of strength J with the conduction electron spin-density $\vec{s}_0 = \sum_{\sigma\sigma'} f^+_{0\sigma} \vec{\sigma}_{\sigma\sigma'} f_{0\sigma'}$ at the impurity site, where $f_{0\sigma} = \sum_k c_{k\sigma}$ the local Wannier state of the conduction electrons at the impurity site. The connection to the Anderson model can be established formally via a Schrieffer-Wolff transformation. In essence, provided $\varepsilon_d < 0$ and $\varepsilon_d + U > 0$ so that a single electron occupies the local level in the Anderson model, the physics of both models will be the same at low temperatures.[4] In this case, one finds the correspondence $J = 2V^2(\frac{1}{U+\varepsilon_d} - \frac{1}{\varepsilon_d})$, which reduces to $8V^2/U$ for the symmetric case $\varepsilon_d = -U/2$ (see discussion of zero bandwidth limit below).
For a numerical treatment of the Anderson and Kondo models, it is useful to reformulate them in the form of linear chain models [2, 3]. This will allow them to be iteratively diagonalized by a procedure to be described in Sec. 3. We carry this out for the Anderson model: first notice that the impurity state in the Anderson model hybridizes with a local Wannier state $|0\sigma\rangle = f^+_{0\sigma}|{\rm vac}\rangle$, with $|{\rm vac}\rangle$ the vacuum state, and $f^+_{0\sigma}$ given by

$$V f^+_{0\sigma} = \sum_k V_{kd} c^+_{k\sigma}. \tag{6}$$

The value of V follows from the normalization $\{f_{0\sigma}, f^+_{0\sigma}\} = 1$

$$V = (\sum_k |V_{kd}|^2)^{1/2}. \tag{7}$$

Using the above local state one can apply the Lanczos procedure (Appendix B) for tridiagonalizing a Hermitian operator, such as $H_{\rm bath}$, to obtain

$$H_{\rm bath} = \sum_{k\sigma} \varepsilon_k c^+_{k\sigma} c_{k\sigma} \rightarrow \sum_{\sigma,n=0}^{\infty} [\epsilon_n f^+_{n\sigma} f_{n\sigma} + t_n (f^+_{n\sigma} f_{n+1\sigma} + H.c.)] \tag{8}$$

[4] Strictly speaking, one should also include a potential scattering term in the Kondo model, of the form $\sum_{kk'\sigma} V^{\rm pot}_{kk'} c^\dagger_{k\sigma} c_{k'\sigma}$ for this to be true.

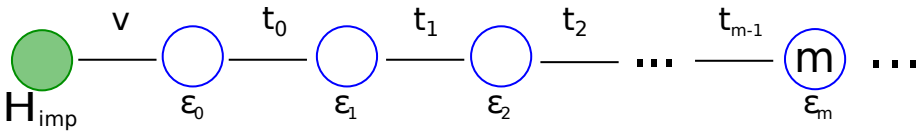


Fig. 3: *The linear chain form of the Anderson model (9).* $H_{\rm imp} = \varepsilon_d n_d + U n_{d\uparrow} n_{d\downarrow}$. *The "site energies"* ε_n *and "hoppings"* t_n *follow from* $\Delta(\omega)$.

with site energies, ϵ_n, and hoppings, t_n, depending only on the dispersion ε_k and hybridization matrix elements V_{kd} through the hybridization function $\Delta(\omega)$ resulting in the linear chain form [2]

$$\begin{aligned} H_{AM} &= \varepsilon_d n_d + U n_{d\uparrow} n_{d\downarrow} + V \sum_\sigma (f^+_{0\sigma} d_\sigma + d^+_\sigma f_{0\sigma}) \\ &+ \sum_{\sigma,n=0}^{\infty} \left[\epsilon_n f^+_{n\sigma} f_{n\sigma} + t_n (f^+_{n\sigma} f_{n+1\sigma} + f^+_{n+1\sigma} f_{n\sigma}) \right] \end{aligned} \tag{9}$$

depicted in Fig. 3 (with $n_d \equiv \sum_\sigma n_{d\sigma}$). Although, formally, this model looks like the one-dimensional real-space models treated by the DMRG method[22, 23], the interpretation here is not in terms of electrons hopping on a one-dimensional lattice in real-space. Instead, as will become clearer in Sec. 3, each successive site added along the chain corresponds to adding lower energy degrees of freedom, measured relative to the Fermi level. By considering longer chains one can then access lower energies.
The same procedure can be used to reformulate any quantum impurity model in terms of an impurity site with local interactions attached to a one-dimensional chain of non-interacting sites. For example, the Kondo model (5) can be rewritten as

$$H_{KM} = -g\mu_B S_z + J\vec{S}.\vec{s}_0 + \sum_{\sigma,n=0}^{\infty} \left[\epsilon_n f^+_{n\sigma} f_{n\sigma} + t_n (f^+_{n\sigma} f_{n+1\sigma} + f^+_{n+1\sigma} f_{n\sigma}) \right] \tag{10}$$

A zeroth order (high energy) approximation to the spectrum of the Anderson model can be obtained by considering just the coupling of the $n = 0$ Wannier state to the impurity and neglecting all others (the zero-bandwidth limit),

$$H_{AM} \approx H_0 \equiv \varepsilon_d n_d + U n_{d\uparrow} n_{d\downarrow} + V \sum_\sigma (f^+_{0\sigma} d_\sigma + d^+_\sigma f_{0\sigma}) \tag{11}$$

There are 16 many-electron states $|n_d, n_0\rangle$, which can be classified by the conserved quantum numbers of total electron number N_{el}, total z-component of spin S^{tot}_z and total spin $\vec{S}$. Using these symmetries we can diagonalize the block matrices $H^0_{N_e,S,S_z}$ to obtain the many-body eigenstates $|N_{el}, S, S_z, r\rangle$ and the corresponding eigenvalues. For example, in the product basis $|n_d\rangle|n_0\rangle$, the Hamiltonian for $N_e = 1, S = 1/2, S_z = \pm 1/2$ is given by

$$H_{N_e=1,S=1/2,S_z=\pm 1/2} = \begin{pmatrix} \varepsilon_d & V \\ V & 0 \end{pmatrix}$$

with eigenvalues

$$E_\pm = (\varepsilon_d \pm \sqrt{\varepsilon_d^2 + 4V^2})/2$$

Proceeding similarly for the other Hilbert spaces (homework), we find that for the particle-hole symmetric case $\varepsilon_d = -U/2$ in the strong correlation limit $U \gg V^2$, the spectrum separates into two groups of states, one group of low energy states lying close to the (singlet) ground state with spacings $\mathcal{O}(V^2/U)$ and one group of high energy states lying at energies $\mathcal{O}(U/2)$ higher and also split by $\mathcal{O}(V^2/U)$. This limit corresponds to a singly occupied impurity level effectively behaving as a $S = 1/2$. In fact, the 8 lowest states correspond to those obtained from a zeroth order approximation to the spectrum of the Kondo model via

$$H_{KM} \approx H_0 \equiv J\vec{S}.\vec{s}_0 = \frac{J}{2}[(\vec{S}+\vec{s}_0)^2 - \vec{S}^2 - \vec{s}_0^2]. \tag{12}$$

The Kondo model is therefore the low energy effective model of the Anderson model in the limit of strong correlations and single occupancy. By comparing the splitting of the two lowest levels in the Kondo model, the singlet and triplet states, with the corresponding splitting of the same levels in the Anderson model one finds the relation between the bare parameters of the models to be $J = 8V^2/U$, in agreement with the value obtained from the Schrieffer-Wolff transformation $J = 2V^2(\frac{1}{U+\varepsilon_d} - \frac{1}{\varepsilon_d})$ upon setting $\varepsilon_d = -U/2$ [24].
Within the above zeroth order approximation, $H \approx H_0$, of the Kondo (and Anderson) model, excitations are unrenormalized. The singlet-triplet excitation (splitting) takes the bare value J. The key ingredient of Wilson's NRG, to be discussed in the next section, is a *controlled procedure* for adding the remaining states $n = 1, 2, \ldots$ neglected in the above zero bandwidth approximation. As we shall see in the calculation of dynamical quantities below, this leads to a drastic renormalization of the spin and single-particle excitations, such that the relevant excitations of the Kondo model are not on the bare scale J but on the Kondo scale $T_{\rm K} = D(\rho J)^{1/2}\exp(-1/\rho J)$, where $\rho = 1/2D$ is the density of conduction states (e.g., see Fig. 7-8 in Sec. 3). One can interpret this large renormalization $J \to T_{\rm K}$ as a renormalization of a bare tunneling amplitude (J) due to the dissipative effects of the bath of conduction electrons by a mapping of the Kondo model onto the dissipative two-state system (also called the spin-boson model). We introduce this briefly in the next subsection, partly to make the above connection, and partly to show that the linear chain representation, which is the starting point for NRG calculations, applies also to bosonic quantum impurity models. For a detailed discussion of the spin-boson model, and its real time dynamics, see lecture B5 by Schoeller.

Spin-boson model

The Hamiltonian of the spin-boson model is given by,

$$H_{\rm SB} = \underbrace{-\frac{1}{2}\Delta_0\sigma_x + \frac{1}{2}\epsilon\sigma_z}_{H_{\rm imp}} + \underbrace{\frac{1}{2}\sigma_z\sum_i \lambda_i(a_i + a_i^\dagger)}_{H_{\rm int}} + \underbrace{\sum_i \omega_i(a_i^\dagger a_i + 1/2)}_{H_{\rm bath}}. \tag{13}$$

The first term $H_{\rm imp}$ describes a two-level system with bias splitting ϵ and tunneling amplitude Δ_0, and $\sigma_{i=x,y,z}$ are Pauli spin matrices. The third term, $H_{\rm bath}$, is the environment and consists of an infinite set of harmonic oscillators ($i = 1, 2, \ldots, \infty$) with $a_i(a_i^\dagger)$ the annihilation (creation) operators for a harmonic oscillator of frequency ω_i and $0 \le \omega_i \le \omega_{\rm c}$, where $\omega_{\rm c}$ is an upper cut-off frequency. The non-interacting density of states of the environment is denoted by $g(\omega_i) = \sum_i \delta(\omega - \omega_i)$ and is finite in the interval $[0, \omega_{\rm c}]$ and zero otherwise. Finally, $H_{\rm int} = \frac{1}{2}\sigma_z\sum_i \lambda_i(a_i + a_i^\dagger)$ describes the coupling of the two-state system co-ordinate

σ_z to the oscillators, with λ_i denoting the coupling strength to oscillator i. The function $\Gamma(\omega+i\delta) = \sum_i (\lambda_i/2)^2/(\omega-\omega_i+i\delta) = \int d\omega' (\lambda(\omega')/2)^2\, g(\omega')/(\omega-\omega'+i\delta)$ characterizes the system-environment interaction. It may be shown, that the Ohmic two-state system , specified by a spectral function $J(\omega) = -\frac{1}{\pi}\mathrm{Im}\Gamma(\omega+i\delta) \sim \alpha\omega$ for $\omega \to 0$, where α is the dimensionless dissipation strength, is equivalent to the anisotropic Kondo model $H = \sum_{k\sigma}\epsilon_k c^\dagger_{k\sigma}c_{k\sigma} + \frac{J_\perp}{2}(S^+ s^-_0 + S^- s^+_0) + J_\parallel S_z s^z_0 + BS_z$, where $J_\perp$ ($J_\parallel$) is the transverse (longitudinal) part of the Kondo exchange interaction and B is a local magnetic field. The correspondence is given by $\rho J_\perp = -\Delta_0/\omega_{\rm c}$ and $\alpha = (1+2\delta/\pi)^2$ where $\delta = \arctan(-\pi\rho J_\parallel/4)$ and ρ is the density of states of the conduction electrons in the anisotropic Kondo model [25, 26, 14, 27, 15]. The low energy scale of the Ohmic two-state system is the renormalized tunneling amplitude $\Delta_{\rm r}$ given by $\Delta_{\rm r}/\omega_{\rm c} = (\Delta_0/\omega_{\rm c})^{1/(1-\alpha)}$ and corresponds to the low energy Kondo scale $T_{\rm K}$ of the anisotropic Kondo model. This connection between the (anisotropic) Kondo and spin-boson models provides another viewpoint on the local dynamics of a Kondo spin in terms of tunneling and dissipation: $T_{\rm K}$ acquires the meaning of a renormalized tunneling amplitude, with frictional effects of the environment leading to a drastic renormalization of $J_\perp \to T_{\rm K}$, particularly in the limit of strong dissipation $\alpha \to 1^-$, corresponding to the isotropic $J \ll 1$ limit of the Kondo model. Indeed, for $\alpha > 1$, the above corresponence states that $J_\parallel < 0$, which corresponds to the ferromagnetic sector of the Kondo model. Since in this limit, $J_\perp$ is irrelevant [19, 24], it follows that $T_{\rm K}$ and hence Δ_r vanish for $\alpha > 1$, i.e., the frictional effects of the environment are so large for $\alpha > 1$ that quantum mechanical tunneling is destroyed at $T = 0$ and tunneling between the two states is possible only via thermal activation. Of interest also, is the transition from quantum coherent dynamics at weak dissipation $\alpha \ll 1$ to incoherent dynamics at strong dissipation $\alpha \to 1^-$, see lecture B3 by Schoeller.

For a numerical treatment using the NRG, one proceeds to re-formulate the model (13) in a linear chain form as in (9) and (10) for the Anderson and Kondo models. Thus, one uses the Lanczos procedure and applies $H_{\rm bath}$ repeatedly on the local bosonic orbital $\lambda b_0 = \sum_i \lambda_i a_i$ to tridiagonalize $H_{\rm bath}$. The resulting linear chain model,

$$H_{\rm SB} = \underbrace{-\frac{1}{2}\Delta_0\sigma_x + \frac{1}{2}\epsilon\sigma_z}_{H_{\rm imp}} + \underbrace{\frac{1}{2}\sigma_z\lambda(b_0+b_0^\dagger)}_{H_{\rm int}} + \underbrace{\sum_{m=0}^{\infty}\epsilon_m b_m^\dagger b_m + t_m(b_m^\dagger b_{m+1} + b_{m+1}^\dagger b_m)}_{H_{\rm bath}}, \qquad (14)$$

may then be treated within NRG in a similar way to the treatment of the Anderson and Kondo models [28]. One difference, is that the number of bosons in the states $b_m^\dagger|vac\rangle$ is arbitrary, requiring an additional approximation even at the first iteration for $H_0 = H_{\rm imp} + \frac{1}{2}\sigma_z\lambda(b_0+b_0^\dagger)$ to restrict the maximum number of bosons to n_b (typically $8-10$).

3 Wilson's Numerical Renormalization Group Method

Wilson's formulation of the RG for the Kondo model is similar in spirit to Anderson's scaling method (see Hewson's book [24] or Ref. [19]). The main difference lies in the non-perturbative construction of the RG transformation using a numerical representation of the effective Hamiltonians. The scaling approach uses perturbation theory in the initially small dimensionless coupling (J/D) to construct such a transformation, but since J/D increases with decreasing energy scale this approach eventually becomes inaccurate. In the Wilson approach the RG transformation is perturbative only via a small parameter $\Lambda^{-1/2} < 1$ which is related to the momentum

rescaling factor $\Lambda > 1$. The accuracy of the transformation is the same at each step and is independent of the size of the running couplings. For this reason it gave the first correct description of the crossover from the weak coupling to the strong coupling regime of the Kondo model. The NRG procedure, described in detail below, involves three steps, illustrated in Fig. 4 (a)-(c) for the Anderson model.

Separation of scales and logarithmic discretization approximation

In the Kondo problem, as in other quantum impurity problems, the behaviour of the system changes qualitatively over many energy scales as it passes through a crossover between fixed points (e.g. from behaviour characteristic of a well defined magnetic moment at high temperature to behaviour characteristic of a Fermi liquid at temperatures below the crossover scale). In order to describe this crossover the idea is to separate out the many energy scales in the problem, which arise from the conduction band $[-D, +D]$, and to set up a procedure for treating each scale in turn. We saw in the previous section that it is always possible to rewrite a quantum impurity model in the form of a (semi-infinite) linear chain, see Fig. 3. Truncating this chain to include orbitals $n = 0, \ldots, m$, we have

$$\begin{aligned} H_{AM} &\approx H_m = \varepsilon_d n_d + U n_{d\uparrow} n_{d\downarrow} + V \sum_\sigma (f^+_{0\sigma} d_\sigma + d^+_\sigma f_{0\sigma}) \\ &+ \sum_{\sigma,n=0}^{m} \epsilon_n f^+_{n\sigma} f_{n\sigma} + \sum_{\sigma,n=0}^{m-1} t_n (f^+_{n\sigma} f_{n+1\sigma} + f^+_{n+1\sigma} f_{n\sigma}) \end{aligned} \tag{15}$$

with the truncated Hamiltonians H_m satisfying the recursion relation

$$H_{m+1} = H_m + \sum_\sigma \epsilon_m f^\dagger_{m\sigma} f_{m\sigma} + t_m \sum_\sigma (f^+_{m\sigma} f_{m+1\sigma} + f^+_{m+1\sigma} f_{m\sigma}). \tag{16}$$

Hence, it appears that with this recursion relation, one can iteratively diagonalize the Anderson model starting from H_0, which we diagonalized explicitly in the previous section. At some point, we will need to retain only the lowest many-body states of H_m since the Hilbert space grows as $4 \times 4^{m+1}$. The validity of this procedure then depends on whether the perturbation in Eq. 16, the last term involving t_m, is small, once we start neglecting some higher energy states.[5] In practice, for a quasi-continuous band $H_{\rm bath} = \sum_{k\sigma} c^\dagger_{k\sigma} c_{k\sigma} = \sum_\sigma \int_{-D}^{+D} \epsilon c^\dagger_{\epsilon\sigma} c_{\epsilon\sigma}$ the hoppings t_m do not decay with increasing m, and the above procedure breaks down after some iterations. For example, it can be easily shown that for a semi-elliptic density of states $\rho(\epsilon) = \frac{2}{\pi D^2}\sqrt{D^2 - \epsilon^2}$, that $t_m = D/2$ for all m (see Hewson's book[24]).

In order to have a working procedure, involving decreasing hoppings t_m along the chain, and at the same time achieve the energy scale separation described above, Wilson discretized the conduction band into positive and negative energy intervals, $D^+_n = [\Lambda^{-(n+1)}, \Lambda^{-n}]$ and $D^-_n = [-\Lambda^{-n}, -\Lambda^{-(n+1)}]$, $n = 0, 1, \ldots$, about the Fermi level $\epsilon_F = 0$ as shown in Fig. 4 a. The quasi-continuous band was then approximated by a discrete one by keeping only a single conduction state from each interval $D^\pm_n$,

$$H_{\rm bath} = \sum_\sigma \int_{-D}^{+D} \epsilon c^\dagger_{\epsilon\sigma} c_{\epsilon\sigma} \approx \sum_{n=0}^{\infty} (\varepsilon_{-n} c^+_{-n\sigma} c_{-n\sigma} + \varepsilon_{+n} c^+_{+n\sigma} c_{+n\sigma}) \tag{17}$$

[5] If we keep extending the system by one orbital at a time without neglecting any states, no error is made. The onsite term in ϵ_m is diagonal and shifts the low energy levels of H_m.

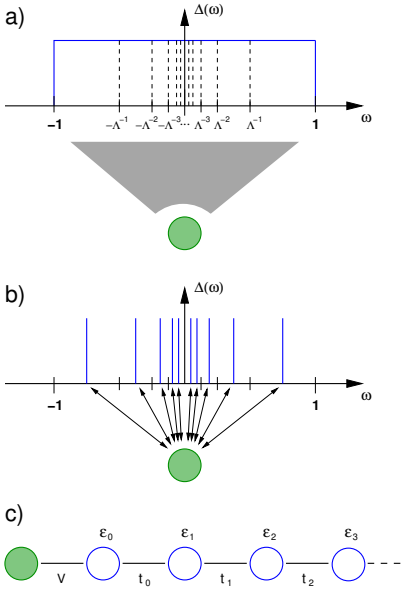


Fig. 4: *Steps in the NRG procedure for an Anderson impurity (circle) coupled to a continuum of conduction states via a hybridization function $\Delta(\omega)$. (a) Logarithmic discretization of the continuum conduction band about the Fermi level $\epsilon_F = 0$ into discrete intervals $D_n^+ = [\Lambda^{-(n+1)}, \Lambda^{-n}]$ and $D_n^- = [-\Lambda^{-n}, -\Lambda^{-(n+1)}]$, $n = 0, 1, \ldots$. (b) Within each discrete interval, choose the conduction state most localized on the impurity. (c) Transform logarithmically discretized model to linear chain form, with hoppings now decreasing along the chain and iteratively diagonalize.*

with

$$\varepsilon_{\pm n} = \pm\frac{1}{2}D(\Lambda^{-n} + \Lambda^{-(n+1)}) = \pm\frac{1}{2}D\Lambda^{-n}(1+\Lambda^{-1}) \tag{18}$$

The states $c^\dagger_{\pm n\sigma}|vac\rangle$ appearing above are the states in each interval $D^\pm_{-n}$ which are most localized near the impurity [2], while the neglected states being orthogonal to these have their wavefunctions localized away from the impurity and are consequently less important (for a more detailed derivation and justification of the logarithmic discretization approximation see Appendix A). By formulating the Anderson model as a linear chain using the above logarithmically discretized conduction band, we obtain the same equations (15-16) as above, but, crucially, with hopping parameters t_m (and onsite energies ϵ_m) that now decay exponentially along the chain. For example, for a constant density of states $\rho(\omega) = 1/2D$ and constant hybridization function $\Delta(\omega) = \pi\rho(\omega)V^2 = \Delta_0$ and $\epsilon_F = 0$ one finds for $m = 0, 1, \ldots$ [1]

$$\epsilon_m = 0, \tag{19}$$

$$t_m = \frac{1}{2}D(1+\Lambda^{-1})\Lambda^{-m/2}\xi_m, \tag{20}$$

$$\xi_m = \frac{1-\Lambda^{-m-1}}{\sqrt{(1-\Lambda^{-2m-1})(1-\Lambda^{-2m-3})}}. \tag{21}$$

The ξ_m converge rapidly to 1 with increasing m and we may write $t_m \approx \frac{1}{2}D(1+\Lambda^{-1})\Lambda^{-m/2}$, so that the Anderson model in Eq. (9) becomes

$$H_{AM} = \varepsilon_d n_d + U n_{d\uparrow} n_{d\downarrow} + V\sum_\sigma (f^+_{0\sigma} d_\sigma + d^+_\sigma f_{0\sigma})$$
$$+ \frac{1}{2}D(1+\Lambda^{-1})\sum_{\sigma,n=0}^{\infty}\Lambda^{-n/2}(f^+_{n\sigma}f_{n+1\sigma} + f^+_{n+1\sigma}f_{n\sigma}) \tag{22}$$

This Hamiltonian provides a clear separation of the energy scales $\frac{1}{2}(1+\Lambda^{-1})\Lambda^{-n/2}, n = 1, 2, \ldots$ in H and allows the diagonalization of the Hamiltonian in a sequence of controlled steps, each

step corresponding to adding an orbital $f_{n\sigma}$ which is a relative perturbation of strength $\Lambda^{-1/2} < 1$, thereby ensuring convergence of the method. This procedure is described in the following two subsections, where we henceforth restrict ourselves to a constant hybridization with hoppings $t_m \approx \frac{1}{2}D(1+\Lambda^{-1})\Lambda^{-m/2}$. The procedure is easily generalized to any hybridization function $\Delta(\omega)$ with hoppings t_m decaying sufficiently fast along the chain.

Renormalization Group Transformation

A renormalization group transformation relating effective Hamiltonians on successive energy scales $\Lambda^{-n/2}$ and $\Lambda^{-(n+1)/2}$ can be set up as follows. First, H in (22) is truncated to H_m, whose lowest scale is $D_m = \frac{1}{2}D(1+\Lambda^{-1})\Lambda^{-(m-1)/2}$. In order to look for fixed points we define rescaled Hamiltonians $\bar{H}_m \equiv H_m/D_m$ such that the lowest energy scale of $\bar{H}_m$ is always of $\mathcal{O}(1)$:

$$\begin{aligned}\bar{H}_m &= \Lambda^{(m-1)/2}[\sum_{n=0}^{m-1}\Lambda^{-n/2}(f^+_{n\sigma}f_{n+1\sigma}+f^+_{n+1\sigma}f_{n\sigma}) \\ &+ \tilde{\varepsilon}_d n_d + \tilde{U}n_{d\uparrow}n_{d\downarrow} + \tilde{V}\sum_\sigma(f^+_{0\sigma}d_\sigma + d^+_\sigma f_{0\sigma})], \qquad (23)\\ \tilde{\varepsilon}_d &= \frac{\varepsilon_d}{\frac{1}{2}D(1+\Lambda^{-1})},\ \tilde{V}=\frac{V}{\frac{1}{2}D(1+\Lambda^{-1})},\ \tilde{U}=\frac{U}{\frac{1}{2}D(1+\Lambda^{-1})}, \qquad (24)\end{aligned}$$

from which we can recover H as

$$H = \lim_{m\to\infty}\frac{1}{2}D(1+\Lambda^{-1})\Lambda^{-(m-1)/2}\bar{H}_m. \qquad (25)$$

The sequence of rescaled Hamiltonians $\bar{H}_m$ satisfies the recursion relation

$$\bar{H}_{m+1} = \Lambda^{1/2}\bar{H}_m + (f^+_{m\sigma}f_{m+1\sigma} + f^+_{m+1\sigma}f_{m\sigma}), \qquad (26)$$

and allows a RG transformation $\mathcal{T}$ to be defined:

$$\bar{H}_{m+1} = \mathcal{T}[\bar{H}_m] \equiv \Lambda^{1/2}\bar{H}_m + (f^+_{m\sigma}f_{m+1\sigma} + f^+_{m+1\sigma}f_{m\sigma}) - \bar{E}_{G,m+1} \qquad (27)$$

with $\bar{E}_{G,m+1}$ the ground state energy of $\bar{H}_{m+1}$. In fact $\mathcal{T}$ defined in (27) does not have fixed points since it relates a Hamiltonian with an even number of orbitals to a Hamiltonian with an odd number of orbitals. The even/odd spectra do not match for the Kondo model. However, $\mathcal{R} = \mathcal{T}^2$, can be defined as the RG transformation and this will have fixed points, a set of even m fixed points and a set of odd m fixed points:

$$\bar{H}_{m+2} = \mathcal{R}[\bar{H}_m] \equiv \mathcal{T}^2[\bar{H}_m] \qquad (28)$$

Iterative diagonalization scheme

The transformation R relates effective Hamiltonians $H_m = D_m\bar{H}_m$ and $H_{m+1} = D_{m+1}\bar{H}_{m+1}$ on decreasing scales $D_m > D_{m+1}$. It can be used to iteratively diagonalize the Anderson Hamiltonian by the following sequence of steps:

1. the local part

$$\bar{H}_0 = \Lambda^{-1/2} \left[\tilde{\varepsilon}_d n_d + \tilde{U} n_{d\uparrow} n_{d\downarrow} + \tilde{V} \sum_\sigma (f^+_{0\sigma} d_\sigma + d^+_\sigma f_{0\sigma}) \right], \tag{29}$$

which contains the many-body interactions, is diagonalized (the "zeroth" order step described in Sec. 2),

2. assuming that $\bar{H}_m$ has been diagonalized for some $m \geq 0$,

$$\bar{H}_m = \sum_q \bar{E}^m_q |q\rangle\langle q| \tag{30}$$

we add a "site" and use (27) to set up the matrix for $\bar{H}_{m+1}$ within a product basis

$$|q, \alpha_{m+1}\rangle = |q\rangle_m |\alpha_{m+1}\rangle \tag{31}$$

consisting of the eigenstates $|q\rangle_m$ of $\bar{H}_m$ and the 4 states $|\alpha_{m+1}\rangle$ of the next orbital along the chain (i.e. $|0\rangle, |\uparrow\rangle \equiv f^\dagger_{m+1\uparrow}|0\rangle, |\downarrow\rangle \equiv f^\dagger_{m+1\downarrow}|0\rangle$, and $|\uparrow\downarrow\rangle = f^\dagger_{m+1\uparrow} f^\dagger_{m+1\downarrow}|0\rangle$). The resulting matrix

$$\begin{aligned} \langle q, \alpha_{m+1}|\bar{H}_{m+1}|q', \alpha'_{m+1}\rangle &= \Lambda^{1/2} \delta_{\alpha_{m+1},\alpha'_{m+1}} \delta_{q,q'} \bar{E}^m_q \\ &+ (-1)^{N_{e,q'}} {}_m\langle q|f^+_{m\sigma}|q'\rangle_m \langle \alpha_{m+1}|f_{m+1\sigma}|\alpha'_{m+1}\rangle \\ &+ (-1)^{N_{e,q}} \langle \alpha_{m+1}|f^+_{m+1\sigma}|\alpha'_{m+1}\rangle_m \langle q|f_{m\sigma}|q'\rangle_m, \end{aligned} \tag{32}$$

with $N_{e,q}, N_{e,q'}$ the number of electrons in $|q\rangle, |q'\rangle$ respectively, is diagonalized and the procedure is repeated for the next energy shell as depicted in Fig. 4 c. Since $\bar{H}_m$ is already diagonalized, the off-diagonal matrix elements, involving ${}_m\langle q|f^\dagger_{m\sigma}|q'\rangle_m = {}_m\langle q'|f_{m\sigma}|q\rangle^\dagger_m$, can be expressed in terms of the known eigenstates of $\bar{H}_m$ by using the unitary transformation relating the product states $|q\rangle_{m-1}|\alpha_m\rangle$ to the eigenstates $|q\rangle_m$ of $\bar{H}_m$,

$$|q\rangle_m = \sum_{r,\alpha_m} U_m(r\alpha_m, q)|q\rangle_{m-1}|\alpha_m\rangle, \tag{33}$$

where U_m is the matrix of eigenvectors of $\bar{H}_m$. (for explicit expressions of the resulting off-diagonal terms in Eq. (32) see [2]).

Equation (33) also shows that the NRG eigenstates have the form of so called matrix product states (MPS) [29], a feature of NRG shared also by the density matrix renormalization group method (DMRG) for one-dimensional quantum systems (see Appendix D for a comparison of NRG with this method). In order to see this, we introduce the notation $A^{\alpha_m}_{q_m q_{m-1}} \equiv U_m(q_{m-1}\alpha_m, q_m)$ with $|q_m\rangle \equiv |q\rangle_m$ and repeatedly apply Eq. (33) to obtain

$$\begin{aligned} |q_m\rangle &= \sum_{q_{m-1},\alpha_m} A^{\alpha_m}_{q_{m-1}q_m} |q_{m-1}\rangle|\alpha_m\rangle \\ &= \sum_{q_{m-1},\alpha_m} A^{\alpha_m}_{q_{m-1}q_m} \left[\sum_{q_{m-2}\alpha_{m-1}} A^{\alpha_{m-1}}_{q_{m-2}q_{m-1}} |\alpha_{m-1}\rangle \right] |\alpha_m\rangle \\ &= \sum_{q_{m_0},\alpha_{m_0+1}\ldots\alpha_m} \left(A^{\alpha_{m_0+1}} \ldots A^{\alpha_m}\right)_{q_{m_0}q_m} |\alpha_{m_0+1}\rangle \ldots |\alpha_m\rangle, \end{aligned} \tag{34}$$

where $m_0 \geq 0$.

3. In order to reduce the size of the matrices that need to be diagonalized, one uses available symmetries, such as conservation of charge (electron number relative to half filling) $[Q_m, H_m] = 0$, with $Q_m = (n_d - 1) + \sum_{n=0}^{m} \sum_{\sigma} (f_{n\sigma}^{\dagger} f_{n\sigma} - 1)$, or conservation of total spin $[\vec{S}_m^2, H_m] = 0$. For multi-channel models, such as the three channel Kondo model, additional symmetries, such as $SU(3)$, may be used to significantly reduce the numerical effort [30, 31]. The use of symmetries, beyond the advantage of reducing the computational cost, also improves the accuracy of the calculations once one starts to neglect high energy states (see next subsection), since it avoids the possibility of splitting up degenerate states within a multiplet carrying the same conserved quantum numbers.

Truncation

In practice since the number of many-body states in $\bar{H}_m$ grows as $4 \times 4^{m+1}$ it is not possible to retain all states after a given iteration $m = m_0$, where m_0 is typically 4 or 5 for single-channel models. For $m \geq m_0$ only the lowest 1000 or so states of $\bar{H}_m$ are retained, we call these the kept states and denote them by $|k\rangle_m$, while the higher energy states neglected are denoted by $|l\rangle_m$, see Fig. 11 in Sec. 5. While only kept states are used to set up and diagonalize the sequence of Hamiltonians $\bar{H}_m, m = m_0, m_0 + 1, ...$ up to a maximum chain size of length $m = N$, we shall see later that the discarded states $|l\rangle_m$ from each iteration $m \geq m_0$ prove to be very useful for calculating physical properties.
The truncation of the spectrum of $\bar{H}_m$ restricts the range of eigenvalues in $H_m = D_m \bar{H}_m$ to be such that $0 \leq E_q^m \leq K D_m$ where $K = K(\Lambda)$ depends on Λ and the number of states retained. For 1000 states and $\Lambda = 3$, $K(\Lambda) \approx 10$. However, eigenvalues below D_m are only approximate eigenvalues of the infinite system H, since states with energies below D_m are calculated more accurately in subsequent iterations $m+1, m+2, \ldots$. Therefore the part of the spectrum of H_m which is close to the spectrum of H is restricted to $D_m \leq E_q^m \leq K(\Lambda) D_m$. This allows the whole spectrum of H to be recovered by considering the spectra of the sequence of Hamiltonians H_m, $m = 0, 1, \ldots$. In this way the many–body eigenvalues and eigenstates are obtained on all energy scales. Due to the smallness of the perturbation (of $\mathcal{O}(\Lambda^{-1/2}) < 1$) in adding an energy shell to go from H_m to H_{m+1}, the truncation of the high energy states turns out, in practice, to be a very good approximation.

Fixed Points

The analysis of fixed points is important to gain a conceptual understanding of the model and for accurate analytic calculations in the vicinity a fixed point [2].
From (28), a fixed point H^* of $\mathcal{R} = \mathcal{T}^2$ is defined by

$$H^* = \mathcal{R}[H^*]. \tag{35}$$

Proximity to a fixed point is identified by ranges of m, $m_1 \leq m \leq m_2$, where the energy levels $\bar{E}_p^m$ of $\bar{H}_m$ are approximately independent of m: $\bar{E}_p^m \approx \bar{E}_p$ for $m_1 \leq m \leq m_2$. A typical energy level flow diagram showing regions of m where the energy levels are approximately constant is shown in Fig. 5 (a) for the anisotropic Kondo model (AKM) [26]:

$$H_{AKM} = \sum_{k\sigma} \varepsilon_k c_{k\sigma}^{+} c_{k\sigma} + \frac{J_{\perp}}{2} (S^{+} f_{0\downarrow}^{+} f_{0\uparrow} + S^{-} f_{0\uparrow}^{+} f_{0\downarrow}) + \frac{J_{\parallel}}{2} S_z (f_{0\uparrow}^{+} f_{0\uparrow} - f_{0\downarrow}^{+} f_{0\downarrow}) \tag{36}$$

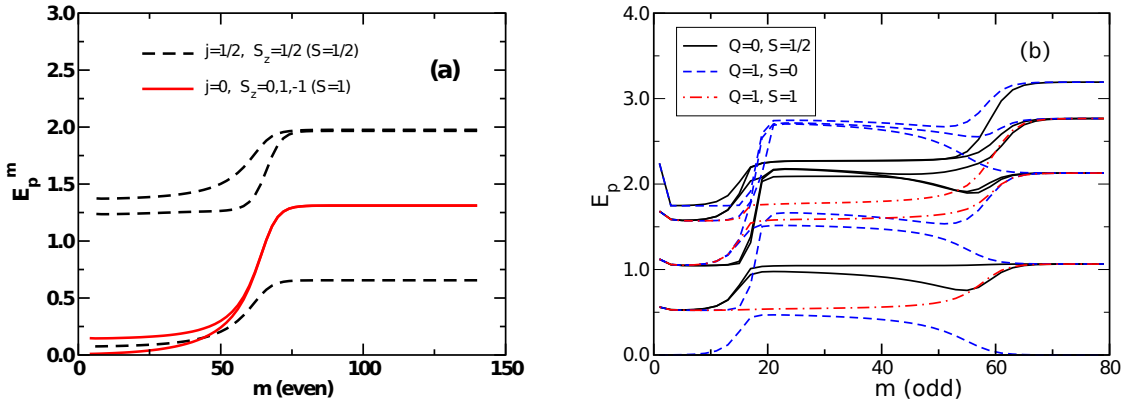


Fig. 5: (**a**) *The lowest rescaled energy levels of the AKM for* $J_{\|} = 0.443$ *and* $J_{\perp} = 0.01$. *The states are labeled by conserved pseudospin* j *and total* S_z *(adapted from [26]).* (**b**) *The flow of the lowest many-body energy levels of the Anderson model for* $\varepsilon_d = -U/2$, $U/\pi\Delta_0 = 12.66$, *and* $\Lambda = 2.5$. *States are labeled by quantum numbers for total charge* Q *and total spin* S *(adapted from [4, 2]).*

There is an unstable high energy fixed point (small m) and a stable low energy fixed point (large m). The low energy spectrum is identical to that of the isotropic Kondo model at the strong coupling fixed point $J = \infty$ in [1] (e.g. the lowest single particle excitations in Fig. 5 (a), $\eta_1 = 0.6555$, $\eta_2 = 1.976$ agree with the $\Lambda = 2$ results of the isotropic model in [1]). The crossover from the high energy to the low energy fixed point is associated with the Kondo scale $T_{\rm K}$. Spin-rotational invariance, broken at high energies, is restored below this scale [e.g., the $j = 0$ states with $S_z = 0$ and $S_z = \pm 1$ become degenerate below $T_{\rm K}$ and can be classified by the same total spin S as indicated in Fig. 5 (a)]. In Fig. 5 (b) typical energy level flows for the symmetric Anderson impurity model $\varepsilon_d = -U/2$ in the strongly correlated Kondo regime are shown. Here, one sees three fixed points: an unstable free orbital fixed point for $m < 10$, a marginal fixed point for $10 < m < 50$ corresponding to formation of a local moment interacting weakly via the antiferromagnetic Kondo exchange with the conduction electrons. In this region, the effective Hamiltonian is essentially the Kondo model. Finally, for $m > 50$ there is a flow to the stable strong coupling fixed point, characterized by a fixed point spectrum obtained by setting $J = \infty$, i.e. the local spin and local conduction orbital are frozen out. The fixed point spectrum is then that of a free electron chain with one site removed, i.e., there is a crossover to an even m fixed point spectrum. The freezing out of the local spin, implies that inelastic scattering processes are blocked as $m \to \infty$ ($T \to 0$) and one obtains the picture of a renormalized Fermi liquid at low temperatures.

Analytic calculations can be carried out in the vicinity of these various fixed points by setting up effective Hamiltonians $H_{\rm eff} = H^* + \sum_\lambda \omega_\lambda O_\lambda$, where the leading deviations O_λ about H^* can be obtained from general symmetry arguments. This allows, for example, thermodynamic properties to be calculated in a restricted range of temperatures, corresponding to the restricted range of m where $\bar{H}_m$ can be described by a simple effective Hamiltonian $H_{\rm eff}$. In this way Wil-

son could show that the ratio of the impurity susceptibility, $\chi_{\rm imp}$, and the impurity contribution to the linear coefficient of specific heat, $\gamma_{\rm imp}$, at $T = 0$, is twice the value of a non-interacting Fermi liquid: $R = \frac{4\pi^2\chi_{imp}}{3\gamma_{imp}} = 2$. We refer the reader to the detailed description of such calculations in [1, 2], and we turn now to the numerical procedure for calculating thermodynamics, which can give results at all temperatures, including the crossover regions.

4 Calculation of physical properties

The ability of the method to yield thermodynamic, dynamic and transport properties makes it very useful for interpreting experimental results [6]. We shall first describe the calculation of thermodynamics and dynamics using conventional approaches (without use of the complete basis set, but inlcuding reduced density matrices for dynamics) [2, 7, 8]. In Sec. 5 we shall then discuss more recent approaches using the complete basis set and full density matrix [33, 10] (this division, however, is somewhat arbitrary).

Thermodynamics: conventional approach

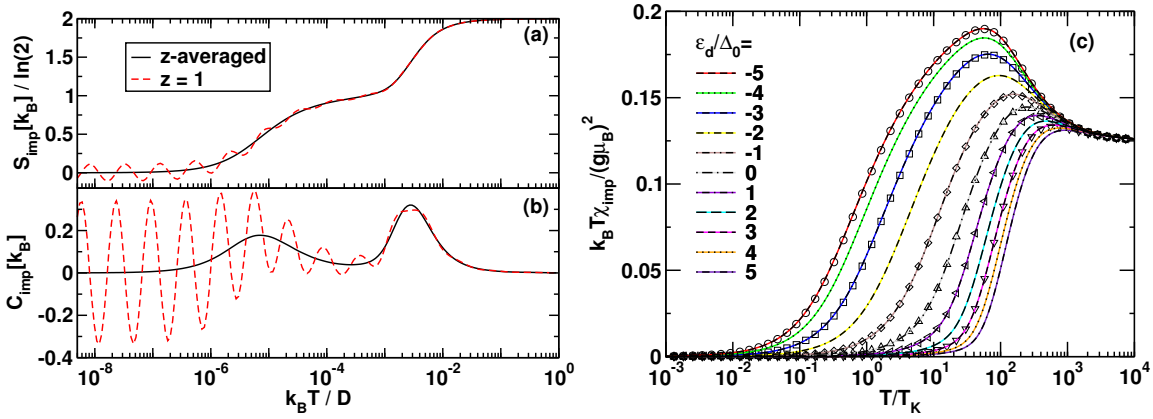


Fig. 6: *Temperature dependence of, (a), the impurity entropy, $S_{\rm imp}(T)$, and, (b), the impurity specific heat, $C_{\rm imp}(T)$, for the symmetric Anderson model with $U/\Delta_0 = 12$ and $\Delta_0 = 0.001D$. The calculations are for $\Lambda = 4$, without z-averaging [$n_z = 1$, $z = 1$ (dashed lines)], and with z-averaging [$n_z = 2$, $z = 1/4,\ 3/4$ (solid lines)]. For $\Lambda = 4$ two z values suffice to eliminate the discretization oscillations[34]. (c) Impurity susceptibility, $\chi_{\rm imp}(T)$, vs $T/T_{\rm K}$ for the asymmetric Anderson model with same parameters as above and several values of ε_d/Δ_0 with $T_{\rm K}$ for symmetric model. Broken lines: FDM approach. Solid lines: conventional approach. Symbols: Bethe ansatz (for selected values of $\varepsilon_d/\Delta_0 = -5, -3, -1, 0, +1, +3$). NRG parameters: $\Lambda = 10$ with z-averaging [$n_z = 4$, $z = 1/8,\ 1/2,\ 3/8,\ 3/4$][35]*

Suppose we have diagonalized exactly the Hamiltonian for a quantum impurity model such as the Kondo model and that we have all the many-body eigenvalues E_q and eigenstates $|q\rangle$:

$$H = \sum_q E_q |q\rangle\langle q| \equiv \sum_q E_q X_{qq}. \tag{37}$$

[6] Spatial correlations may also be investigated, see Ref. [32]

We can then calculate the partition function

$$Z(T) \equiv Tr\ e^{-H/k_BT} = \sum_q e^{-E_q/k_BT}, \tag{38}$$

and hence the thermodynamics via the impurity contribution to the free energy $F_{\rm imp}(T) = -k_BT \ln Z/Z_c$, where $Z_c = Tr\ e^{-H_c/k_BT}$ is the the partition function for the non-interacting conduction electrons. In the NRG procedure we can only calculate the "shell partition functions" Z_m for the sequence of truncated Hamiltonians H_m:

$$Z_m(T) \equiv Tr\ e^{-H_m/k_BT} = \sum_q e^{-E_q^m/k_BT} = \sum_q e^{-D_m\bar{E}_q^m/k_BT} \tag{39}$$

We will have $Z_m(T) \approx Z(T)$ provided

1. we choose $k_BT = k_BT_m \ll E^m_{max} = D_mK(\Lambda)$ so that the contribution to the partition function from excited states $E_q^m > D_mK(\Lambda)$, not contained in Z_m, is negligible, and

2. the truncation error made in replacing H by H_m in equating (38) and (39) is small. This error has been estimated in [2] to be approximately $\Lambda^{-1}D_m/k_BT_m$.

Combining these two conditions requires that

$$\frac{1}{\Lambda} \ll \frac{k_BT_m}{D_m} \ll K(\Lambda). \tag{40}$$

The choice $k_BT = k_BT_m \approx D_m$ is reasonable and allows the thermodynamics to be calculated at a sequence of decreasing temperatures $k_BT_m \sim D_m, N = 0, 1, \ldots$ from the truncated partition functions Z_m. The procedure yields essentially exact results. For small $\Lambda \lesssim 3$, the window for choosing the temperature T_m to satisfy Eq. 40 is small, and typically only one such temperature is used for each shell. For larger $\Lambda \gg 1$ one can use many temperatures $T_m^i, i = 1, ..., n_T$ which satisify the above condition, however, for large $\Lambda = 4 - 10$, discretization oscillations become important[36, 37]. This problem is overcome by averaging the results over several discretizations of the band, i.e. one carries out several calculations with discretizations of the band $\pm D, \pm D\Lambda^{-(1-z_k)}, \pm D\Lambda^{-(2-z_k)}, ...$ and averages the results for several $z_k, k = 1, ..., N_z$. Figure 6 (a-b) illustrates this for the entropy and specific heat of the Anderson model. In this way, the conventional approach can recover essentially exact results for thermodynamics. Fig. 6 (c) shows a comparison for the impurity static susceptibility of the Anderson impurity model

$$\chi_{\rm imp}(T) = \frac{(g\mu_B)^2}{k_BT}\left[\frac{1}{Z}Tr\ (S_z^{tot})^2e^{-H/k_BT} - \frac{1}{Z_c}Tr\ (S_{z,c}^{tot})^2e^{-H_c/k_BT}\right]$$

to both Bethe-ansatz results and results obtained within the more recent full density matrix approach to be described below.

Dynamics: conventional approach without reduced density matrices

We consider now the application of the NRG method to the calculation of dynamic properties of quantum impurity models [38, 6, 7, 8]. For definiteness we consider the Anderson

impurity model and illustrate the procedure for the impurity spectral density $A_{d\sigma}(\omega,T) = -\frac{1}{\pi}\mathrm{Im}G_{d\sigma}(\omega,T)$, with

$$G_{d\sigma}(\omega,T) = \int_{-\infty}^{+\infty} d(t-t')e^{i\omega(t-t')}G_{d\sigma}(t-t') \tag{41}$$

$$G_{d\sigma}(t-t') = -i\theta(t-t')\langle[d_\sigma(t),d_\sigma^+(t')]_+\rangle_\varrho \tag{42}$$

with ϱ the density matrix of the system.
Suppose we have all the many-body eigenstates $|q\rangle$ and eigenvalues E_q of the Anderson impurity Hamiltonian H. Then the density matrix, $\varrho(T)$, of the full system at temperature $k_BT = 1/\beta$ can be written

$$\varrho(T) = \frac{1}{Z(T)}\sum_q e^{-\beta E_q}|q\rangle\langle q|, \tag{43}$$

and the impurity Green function can be written in the Lehmann representation as

$$G_{d\sigma}(\omega,T) = \frac{1}{Z(T)}\sum_{q,q'}|\langle q|d_\sigma|q'\rangle|^2\frac{e^{-E_q/k_BT}+e^{-E_{q'}/k_BT}}{\omega-(E_{q'}-E_q)} \tag{44}$$

and the corresponding impurity spectral density $A_{d\sigma}$ as

$$A_{d\sigma}(\omega,T) = \frac{1}{Z(T)}\sum_{q,q'}|M_{q,q'}|^2(e^{-E_q/k_BT}+e^{-E_{q'}/k_BT})\delta(\omega-(E_{q'}-E_q)) \tag{45}$$

with $M_{q,q'} = \langle q|d_\mu|q'\rangle$.
Consider first the $T = 0$ case ($T > 0$ is described in the next section), then

$$A_{d\sigma}(\omega,T=0) = \frac{1}{Z(0)}\sum_q|M_{q,0}|^2\delta(\omega+(E_q-E_0))+\frac{1}{Z(0)}\sum_{q'}|M_{0,q'}|^2\delta(\omega-(E_{q'}-E_0), \tag{46}$$

with $E_0 = 0$ the ground state energy. In order to evaluate this from the information which we actually obtain from an iterative diagonalization of H, we consider the impurity spectral densities corresponding to the sequence of Hamiltonians H_m, $m = 0,1,\ldots,N$,

$$A_{d\sigma}^m(\omega,T=0) = \frac{1}{Z_m(0)}\sum_q|M_{q,0}^m|^2\delta(\omega+E_q^m)+\frac{1}{Z_m(0)}\sum_{q'}|M_{0,q'}^m|^2\delta(\omega-E_{q'}^m). \tag{47}$$

From the discussion on the spectrum of H_m in the previous section, it follows that the groundstate excitations of H_m which are representative of the infinite system H are those in the range $D_m \le \omega \le K(\Lambda)D_m$. Lower energy excitations and eigenstates are calculated more accurately at subsequent iterations, and higher energy excitations are not contained in H_m due to the elimination of the higher energy states at each m. Hence, for fixed m, we can approximately evaluate the spectral density at a characteristic frequency $\omega \approx \omega_m \equiv k_BT_m$ via

$$A_{d\sigma}(\omega,T=0) \approx A_{d\sigma}^m(\omega,T=0), \quad m = 0,1,\ldots,N. \tag{48}$$

In making this approximation, we are assuming that the matrix elements $M_{0,q'}^m$ of the finite system Hamiltonian are the same as those of the infinite system $M_{0,q'}$. This assumption fails, when the an applied field strongly affects the groundstate and low lying excited states, thereby making also the matrix elements for the finite size system $M_{0,q'}^m$ appreciably different to those of the

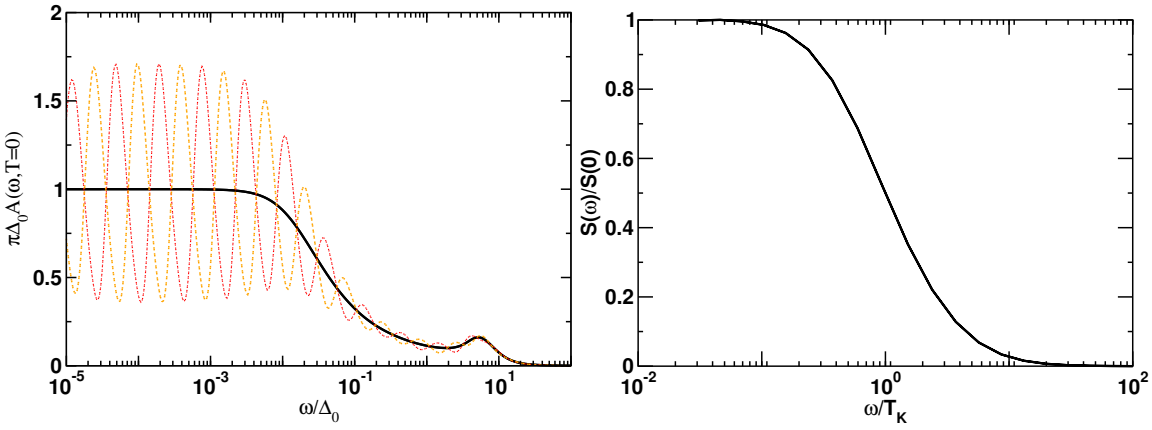


Fig. 7: *(**a**) The impurity spectral density for the symmetric Anderson model for $U/\Delta_0 = 12$ at large $\Lambda = 10$ showing discretization oscillations for two values of z. Averaging over 8 z-values yields the smooth curve. (**b**). Longitudinal spin relaxation function $S(\omega)/S(0)$ versus $\omega/T_{\rm K}$ for the isotropic Kondo model, showing that the spin relaxes incoherently.*

infinite system. We shall come back to this point below, when we introduce the reduced density matrix approach to Green functions[8]. Returning to the calculation of spectral densities, a typical choice for the characteristic frequency to evaluate $A_{d\sigma}(\omega, 0)$ from $A^m_{d\sigma}(\omega, 0)$ is $\omega = 2\omega_m$ for $\Lambda = 2$. In this way $A_{d\sigma}(\omega, T = 0)$ can be calculated at a sequence of decreasing frequencies $\omega = 2\omega_m, m = 0, 1, \ldots, N$ from the quantities $A^m_{d\sigma}$. In practice we are not interested in the discrete spectra $A^m_{d\sigma}(\omega) = \sum_q w^m_q \delta(\omega - E^m_q)$ of the Hamiltonians H_m but in continuous spectra which can be compared with experiment. Smooth spectra can be obtained from the discrete spectra by replacing the delta functions $\delta(\omega - E^m_q)$ by smooth distributions $P_m(\omega - E^m_q)$. A natural choice for the width η_m of P_m is D_m, the characteristic scale for the energy level structure of H_m. Two commonly used choices for P are the Gaussian and the Logarithmic Gaussian distributions [38, 7, 39]. More refined schemes also exist [40, 41] as well as different band discretizations to reduce artifacts close to band edges [42]. A peak of intrinsic width Γ at frequency Ω_0 will be well resolved by the above procedure provided that $\Omega_0 \ll \Gamma$, which is the case for the Kondo resonance and other low energy resonances. In the opposite case, the low (logarithmic) resolution at higher frequencies may be insufficient to resolve the intrinsic widths and heights of such peaks. Usually such higher frequency peaks are due to single-particle processes and can be adequately described by other methods (exceptions include interaction dominated features in the Ohmic two-state system, see below, and in strongly correlated lattice models in high dimensions [43, 16, 44, 45]). In both cases, $\Omega_0 \ll \Gamma$ and $\Omega_0 \gg \Gamma$, the positions and intensities of such peaks is given correctly. An alternative procedure for obtaining smooth spectra, which in principle resolves finite frequency peaks with the same resolution as the low energy peaks, has been proposed in [46] and uses the averaging over several band discretizations, described above for the thermodynamics. This procedure allows carrying out calculations for spectral functions at larger Λ. An example is shown in Fig. 7 (a) for the symmetric Anderson model. As in the thermodynamics, calculations of the dynamics at large $\Lambda \gg 1$ exhibit discretization oscillations, see Fig. 7 (a), which may be eliminated by averaging over several band discretizations.

How accurate is the NRG for dynamic properties ? A good measure of the accuracy of the

procedure is given by the Friedel sum rule, a Fermi liquid relation which states that [24]

$$A_{d\sigma}(0) = \frac{1}{\pi\Delta_0}\sin^2(\pi n_d/2), \qquad n_d = \int_{-\infty}^{0} d\omega\ A_{d\sigma}(\omega) \tag{49}$$

From Fig. 7 (a), we find that $\pi\Delta_0 A_{d\sigma}(0,0) = 1 \pm 10^{-3}$, i.e., the Friedel sum rule is satisfied to within 0.1% relative error. More important, however, is that this error remains small independent of the interaction strength $0 \leq U \leq \infty$. Two-particle Green functions and response functions, such as the longitudinal spin relaxation function,

$$S(\omega) = -\frac{1}{\pi}\frac{\mathrm{Im}\chi_{zz}(\omega)}{\omega}, \qquad \chi_{zz} = \langle\langle S_z; S_z\rangle\rangle,$$

of the Anderson impurity model and of the (anisotropic) Kondo model can also be calculated with comparable accuracy to single-particle spectral functions[26]. The spin relaxation function for the Kondo model is shown in Fig. 7 (b), and illustrates the statement made in Sec. 2, that the spin excitations of the Kondo model are drastically renormalized from the bare value of J down to the Kondo scale T_{K} due to the frictional effects of the environment.
The procedure for calculating finite temperature dynamical quantities, like $A_{d\sigma}(\omega, T)$, required as input for calculating transport properties, is similar to that for $T = 0$ dynamics described above [7]. The spectral density $A_{d\sigma}(\omega, T)$ at fixed temperature T is evaluated as above at frequencies $\omega \approx 2\omega_m, m = 0, 1, \ldots, M \leq N$ until $2\omega_M$ becomes of order k_BT, i.e. $2\omega_M = \alpha k_BT$ with $\alpha \approx 1$. To calculate the spectral density at frequencies $\omega < k_BT$ a smaller "cluster" is used. This is done because when k_BT is larger than the frequency at which the spectral density is being evaluated, it is the excited states of order k_BT contained in previous clusters that are important and not the excitations very much below k_BT. This approach suffers from the same criticism as the $T = 0$ approach above, namely one is using a finite cluster H_m to approximate $M_{q',q} \approx M^m_{q',q}$ (and also $Z(T) \approx Z_m(T)$). In particular, for $\omega < k_BT$, the use of a small cluster of size $M < N$ does not capture the full information available, a definciency which is corrected by the full density matrix approach. Nevertheless, this early approach gives remarkably good results for finite temperature spectra and transport properties[7].

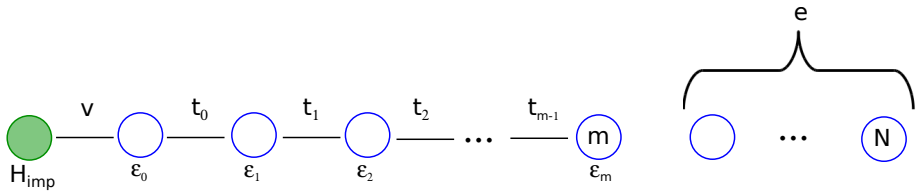


Fig. 8: *The Hilbert space of H_m is supplemented with $N - m$ environment degrees of freedom $e = (\alpha_{m+1}, ..., \alpha_N)$ [8].*

Dynamics: conventional approach with reduced density matrices

A way of reducing finite size errors, inherent in the above approach to Green functions, has been proposed by Hofstetter [8], which we shall describe next. As mentioned above, there are situations when a small field can strongly polarize the low energy states of H_m, thereby strongly affecting the matrix elements $M^m_{q',q}$ and hence the spectra. For example, a magnetic

field $B \approx T_{\rm K}$ in the Anderson model is sufficiently strong to polarize the groundstate such that $n_{d\uparrow} \approx 1$ and $n_{d\downarrow} \approx 0$ at $T \ll T_{\rm K}$. In this case, the use of the canonical density matrix $\varrho(T) \approx \varrho_m(T) = \frac{1}{Z_m(T)} \sum_q e^{-\beta E_q^m} |q\rangle\langle q|$ in evaluating the spectra on scales $\omega_m \gg T_{\rm K}$ can result in large errors. A solution to this is to use $\varrho_N(T) = \frac{1}{Z_N(T)} \sum_q |q\rangle e^{-\beta E_q^N} \langle q|$ for the longest chain diagonalized and to evaluate the Green functions on scales $\omega_m > \omega_N$ by tracing out intermediate degrees of freedom $e = (\alpha_{m+1}, ..., \alpha_N)$ in ϱ_N. Since the longest chain H_N is close to the infinite system limit, this should provide a better description of the spectra, particularly at higher frequencies. In order to carry this procedure, the Hilbert space of each H_m is extended to that of H_N by adding the $N - m$ environment degrees of freedom e, see Fig. 8. Evaluating the reduced density matrix $\varrho_m^{\rm red} = Tr_e\left[\varrho_N\right]$ appearing in Eq. (42) leads to a Lehmann representation for the spectral density at $T = 0$

$$A_{d\sigma}(\omega, T=0) = \sum_{kk'} C_{kk'}^N M_{kk'}^N \delta(\omega - (E_\kappa^N - E_{k'}^N)) \tag{50}$$

$$C_{kk'}^N = \sum_p \varrho_{pk'}^{\rm red} M_{p\kappa}^N + \sum_p \varrho_{kp}^{\rm red} M_{k'p}^N \tag{51}$$

in place of (46). In Fig. 9 (a)-(b), we show a comparison of this approach with results from the

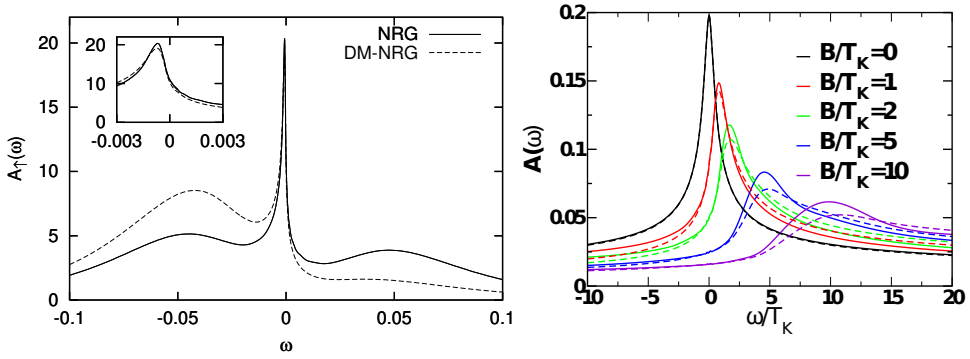


Fig. 9: *Spectral density $A_\uparrow(\omega, T=0)$ for,* (**a**)*, the symmetric Anderson model, with $U/\Delta_0 = 10, \Delta_0 = 0.01$ and $B = \Delta_0/10$, with and without reduced density matrices (DM-NRG/NRG) [8], and,* (**b**)*, for the Kondo model in several magnetic fields B, with and without reduced density matrices (dashed/solid lines, respectively) [12, 47], showing that the low energy Kondo resonance is sufficiently well captured in the conventional approach.*

previous approach for the Anderson and Kondo models in a magnetic field. A field-induced rearrangement of spectral weight at $\omega \approx \varepsilon_d, \varepsilon_d + U$ is well captured by the reduced density matrix approach [Fig. 9 (a)]. The low energy Kondo resonance is less subject to finite size corrections, since this part of the spectrum is already calculated from sufficiently long chains, such that the corrections in using reduced density matrices are small [Fig. 9 (b)].

Transport properties

The transport properties of quantum impurity models, require knowledge of both the frequency and temperature dependence of the impurity spectral density, a topic which was addressed

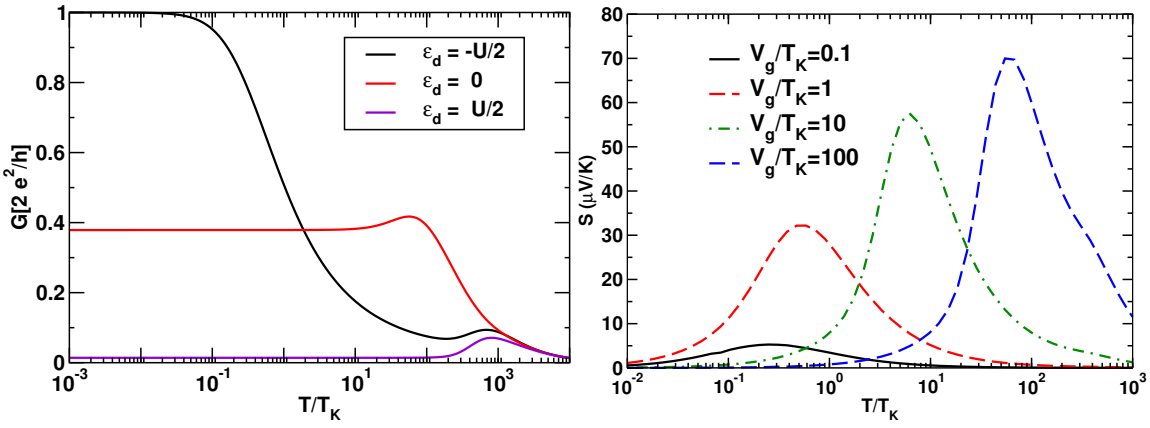


Fig. 10: **(a)** *Linear conductance $G(T)$ versus $T/T_{\rm K}$ for $U/\Delta_0 = 16$ and several values of $\varepsilon_d = -U/2, 0, +U/2$ using the approach of Yoshida et al. [48]. The resistivity of a Kondo impurity $\rho(T)$ is similar to $G(T)$ for the $\varepsilon_d = -U/2$ curve. NRG parameters were for $\Lambda = 4$, $n_z = 2$ [49].* **(b)**. *Thermopower of a negative $U = -16\Delta_0$ quantum dot, exhibiting a large enhancement for gate voltages $V_g \geq T_{\rm K}$[50] (calculated within full density matrix approach to spectral functions).*

above. The linear response conductance, $G(T)$, and thermopower , $S(T)$, through a quantum dot described by the Anderson model, are given by the following expressions (see Lecture B2 by Jakobs for non-equilibrium expressions)

$$G(T) = \frac{e^2}{h} \int d\omega \, \left(-\frac{\partial f}{\partial \omega}\right) \, \sum_\sigma \mathcal{T}_\sigma(\omega, T, B), \tag{52}$$

$$S(T) = -\frac{1}{|e|T} \frac{\int d\omega \, \omega \, (-\partial f/\partial \omega) \sum_\sigma \mathcal{T}_\sigma(\omega)}{\int d\omega \, (-\partial f/\partial \omega) \sum_\sigma \mathcal{T}_\sigma(\omega)}, \tag{53}$$

where the transmission function,$\mathcal{T}_\sigma(\omega, T)$, through a quantum dot symmetrically coupled to left and right leads is related to $A_{d\sigma}(\omega, T)$ via

$$\mathcal{T}_\sigma(\omega, T) = 2\pi\Delta_0 A_{d\sigma}(\omega T).$$

Note, that the discrete form of the spectral function may be directly substituted into the expressions for $G(T)$ and $S(T)$ above, without the necessity of broadening[48]. For the conductance, this leads to

$$G(T) = \frac{\gamma\beta}{Z} \sum_\sigma \sum_{m,n} |M^\sigma_{mn}|^2 \frac{1}{e^{\beta E_m} + e^{\beta E_n}}, \tag{54}$$

with $\gamma = 2\pi\Delta_0 \frac{e^2}{h}$. Results for the temperature dependence of the conductance of the Anderson model, using this procedure, are shown in Fig. 10 (a). Thermoelectric properties have also been investigated for quantum dots with repulsive onsite Coulomb interactions [51] and also for attractive onsite interactions [50]. The latter provide a mechanism for enhancing thermopower, as shown in Fig. 10 (b). The method gives uniformly accurate results at high and low temperatures, as well as correctly describing the crossover region $T \approx T_{\rm K}$ (detailed comparisons of the resistivity of dilute magnetic impurities with known results at high and low temperature can be found [7]). These calculations, and similar resistivity calculations for dilue impurities, provide

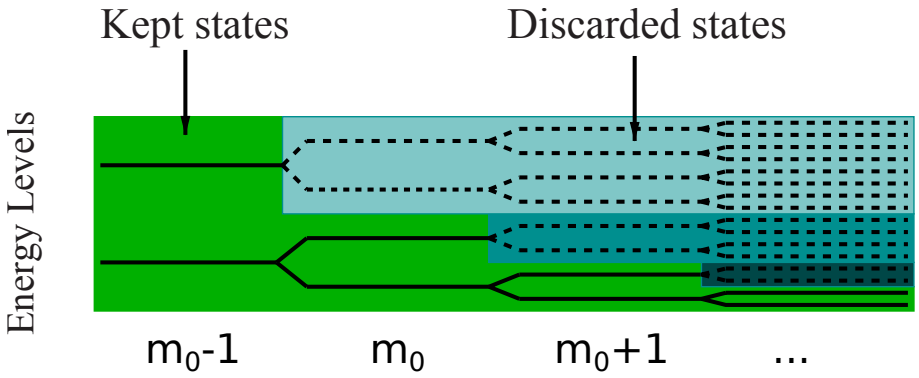


Fig. 11: *For iterations $m < m_0$ all states are kept, while for $m \geq m_0$ only the lowest* 1000 *or so states generated are used to set up the Hamiltonian for the next iteration $m + 1$ (adapted from [54]).*

a quantitative interpretation of experiments for $S = 1/2$ realizations of the Kondo effect . They have also been extended using the full density matrix approach to describe the resistivity and dephasing rates of real Fe impurities in Au and Ag by using a 3-channel Kondo model[52, 31].

5 Complete Basis Set and Full Density Matrix

We noted in Sec. 3, that at each m, the states generated, denoted $|qm\rangle$, are partitioned into the lowest energy retained states, denoted $|km\rangle$, and the high energy eliminated (or discarded) states, $|lm\rangle$. In order to avoid an exponential increase in the dimension of the Hilbert space, only the former are used to set up and diagonalize the Hamiltonian for the next iteration $m + 1$. The eliminated states, while not used in the iterative NRG procedure, may be used to set up a complete orthonormal basis set [33]. This complete basis set is very powerful and allows evaluating correlation functions $\langle A(t)B(0)\rangle$, transient quantities and even thermodynamic expressions in an unambiguous way, avoiding any possible double counting of excitations. Eliminated states from different iterations have no overlap, see Fig. 11, in contrast to the retained states. Hence, using the latter to carry out calculations of physical quantities restricts one to using excitations from a single shell only. However, for finite temperature Green functions and non-equilibrium quantities, multiple-shell contributions become important [53] and the complete basis set offers a way to evaluate these quantities [33].
The complete basis set is defined by the product states $|lem\rangle = |lm\rangle|e\rangle, m = m_0, \ldots, N$, where m_0 is the first iteration at which truncation occurs, and $|e\rangle = |\alpha_{m+1}\rangle|\alpha_{m+2}\rangle \ldots |\alpha_N\rangle$ are environment states at iteration m such that the product states $|lem\rangle$, for each $m = m_0, m_0 + 1, \ldots, N$, reside in the same Fock space (that of the largest system diagonalized, $m = N$). By "e" we shall henceforth denote the collection $e = \{\alpha_{m+1} \ldots \alpha_N\}$. The eliminated states satisfy the completeness relation [33, 55]

$$\sum_{m=m_0}^{N} \sum_{le} |lem\rangle\langle lem| = 1, \tag{55}$$

where for $m = N$ all states are counted as discarded (i.e. there are no kept states at iteration

$m = N$). We shall also use the following representations of this relation, [33, 55]

$$1 = 1_m^- + 1_m^+, \tag{56}$$

$$1_m^- = \sum_{m'=m_0}^{m} \sum_{l'e'} |l'e'm'\rangle\langle l'e'm'| \tag{57}$$

$$1_m^+ = \sum_{m'=m+1}^{N} \sum_{l'e'} |l'e'm'\rangle\langle l'e'm'| = \sum_{ke} |kem\rangle\langle kem|. \tag{58}$$

By using the complete basis set, we can construct the full density matrix FDM [10, 56],

$$\rho = \frac{1}{Z(T)} \sum_{m=m_0}^{N} \sum_{le} e^{-\beta E_l^m} |lem\rangle\langle lem|, \quad \mathrm{Tr}\rho = 1 \Rightarrow \tag{59}$$

$$Z(T) = \sum_{m=m_0}^{N} 4^{N-m} \sum_{l} e^{-\beta E_l^m} \equiv \sum_{m=m_0}^{N} 4^{N-m} Z_m(T) \tag{60}$$

where $Z(T)$ is the partition function made up of the complete spectrum, i.e., it contains the eliminated states from all H_m, $m = m_0, m_0+1, ..., N$. Consequently, it can be used to evaluate the impurity thermodynamics at arbitrary temperatures.
Consider the following density matrix for the m'th shell (defined, however, in the Hilbert space of H_N),

$$\tilde{\rho}_m = \sum_{le} |lem\rangle \frac{e^{-\beta E_l^m}}{\tilde{Z}_m} \langle lem|. \tag{61}$$

Normalization $Tr\,[\tilde{\rho}_m] = 1$ implies that

$$1 = \sum_{l} \frac{e^{-\beta E_l^m}}{\tilde{Z}_m} 4^{N-m} = 4^{N-m} \frac{Z_m}{\tilde{Z}_m} \tag{62}$$

where $Z_m = \sum_l e^{-\beta E_l^m}$. Then the FDM can be written as a sum of weighted density matrices for shells $m = m_0, \ldots, N$

$$\rho = \sum_{m=m_0}^{N} w_m \tilde{\rho}_m \tag{63}$$

$$w_m = 4^{N-m} \frac{Z_m}{Z}; \sum_{m=m_0}^{N} w_m = 1 \tag{64}$$

Application to Thermodynamics

Substituting $\rho = \sum_m w_m \tilde{\rho}_m$ into the expression for the thermodynamic average $\langle \hat{O} \rangle$ of a local observable of the impurity (e.g., n_d or $n_{d\uparrow} n_{d\downarrow}$) and making use of the decomposition of unity Eq. (55), we have

$$\langle \hat{O} \rangle_\rho = \mathrm{Tr}\left[\rho \hat{O}\right] = \sum_{l'e'm'} \langle l'e'm'|\hat{O} \sum_{lem} w_m |lem\rangle \frac{e^{-\beta E_l^m}}{\tilde{Z}_m} \langle lem|l'e'm'\rangle$$

$$= \sum_{lem} O_{ll}^m w_m \frac{e^{-\beta E_l^m}}{\tilde{Z}_m} = \sum_{lm} 4^{N-m} w_m O_{ll}^m \frac{e^{-\beta E_l^m}}{4^{N-m} Z_m} = \sum_{m=m_0,l}^{N} w_m O_{ll}^m \frac{e^{-\beta E_l^m}}{Z_m}, \tag{65}$$

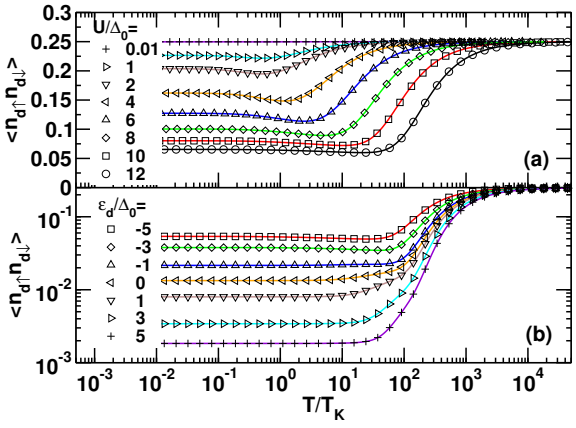


Fig. 12: *Double occupancy* $D_{\rm occ} = \langle n_{d\uparrow} n_{d\downarrow}\rangle$ *as a function of temperature* $T/T_{\rm K}$ *for, (a), the symmetric Anderson model at different* U/Δ_0*, and, (b), for the asymmetric Anderson model at* $U/\Delta_0 = 12$ *and increasing values of* $\varepsilon_d/\Delta_0 = -5, -3, -1, 0, +1, +3, +5$*. FDM (solid lines), conventional approach (symbols).* $\Lambda = 10$ *with* z*-averaging [*$n_z = 4$*,* $z = 1/8,\ 1/2,\ 3/8,\ 3/4$*] [35].*

where orthonormality $\langle lem|l'e'm'\rangle = \delta_{ll'}\delta_{ee'}\delta_{mm'}$, and the trace over the $N-m$ environment degrees of freedom $\sum_{lem}\cdots = \sum_{lm} 4^{N-m}\ldots$ has been used and $O^m_{ll} = \langle lm|\hat{O}|lm\rangle$. For other observables, such as the specific heat, or susceptibility, one requires a similar calculation for the conduction band contribution $\langle\hat{O}\rangle_{\rho_0}$, with ρ_0 the FDM of the non-interacting band. The impurity contribution is then obtained as $O_{\rm imp} = \langle\hat{O}\rangle_{\rho} - \langle\hat{O}\rangle_{\rho_0}$.
For each temperature T and shell m, we require $w_m(T)$ and the factor $B^m_l(T) = e^{-\beta E^m_l}/Z_m$ where $Z_m = \sum_l e^{-\beta E^m_l}$. Numerical problems due to large exponentials are avoided by calculating $B^m_l(T) = e^{-\beta(E^m_l - E^m_0)}/Z'_m$ where $Z'_m = e^{\beta E^m_0} Z_m$ and E^m_0 is the lowest discarded energy for shell m. Figure 12 shows results for the double occupancy of the Anderson model obtained within the FDM approach and comparisons with the conventional approach of Sec. 4.

Application to Dynamics

We consider a general fermionic retarded Green function

$$\begin{aligned} G_{AB}(t) &= -i\theta(t)\langle[A(t),B]_+\rangle \equiv -i\theta(t){\rm Tr}\left[\rho(A(t)B + BA(t))\right] \\ &= -i\theta(t)\left[C_{A(t)B} + C_{BA(t)}\right], \end{aligned} \tag{66}$$

where A, B are fermionic operators, e.g. for the d-level Green function of our quantum dot $A = d_\sigma$ and $B = d^\dagger_\sigma$. The trace is evaluated using the complete basis set. We outline the derivation of $C_{A(t)B}$, with the expression for $C_{BA(t)}$ obtained in a similar manner. We have

$$\begin{aligned} C_{A(t)B} &= {\rm Tr}\left[\rho A(t)B\right] = \sum_{lem}\langle lem|e^{-iHt}Ae^{iHt}B\rho|lem\rangle \\ &= \sum_{lem}\sum_{l'e'm'}\langle lem|e^{-iHt}Ae^{iHt}|l'e'm'\rangle\langle l'e'm'|B\rho|lem\rangle, \end{aligned} \tag{67}$$

which consists of three contributions with $m' = m$, $m' > m$ and $m' < m$. Consider the first contribution ($m' = m$), denoted by $C^{(i)}_{A(t)B}$. Using the NRG approximation $e^{iHt}|l'e'm\rangle \approx e^{iH_m t}|l'e'm\rangle = e^{iE^m_{l'}t}|l'e'm\rangle$ and $\langle lem|A|l'e'm\rangle = \delta_{ee'}\langle lm|A|l'm\rangle = \delta_{ee'}A^m_{ll'}$, we have

$$C^{(i)}_{A(t)B} = \sum_{lm}\sum_{l'} e^{i(E^m_{l'} - E^m_l)t}A^m_{ll'}\sum_e \underbrace{\langle l'em|B\rho|lem\rangle}_{(B\rho)^m_{l'e,le}}$$

Inserting the FDM expression $\rho = \sum_m w_m \tilde{\rho}_m$ into $(B\rho)^m_{l'e,le}$ yields

$$\sum_e (B\rho)^m_{l'e,le} = B^m_{l'l} w_m e^{-\beta E^m_l}/Z_m,$$

hence we have,

$$C^{(i)}_{A(t)B} = \sum_m \frac{w_m}{Z_m} \sum_l \sum_{l'} e^{i(E^m_{l'}-E^m_l)t} A^m_{ll'} B^m_{l'l} e^{-\beta E^m_l}. \tag{68}$$

The off-diagonal contributions with $m' > m$ and $m' < m$ in Eq. (67), which we label by $C^{(ii)}_{A(t)B}$ and $C^{(iii)}_{A(t)B}$, may be put into diagonal form by using $1^+_m = \sum_{m'=m+1}^N \sum_{l'e'} |l'e'm'\rangle\langle l'e'm'| = \sum_{ke} |kem\rangle\langle kem|$ [Eq. (58)], thereby introducing kept states at iteration m (or m') in place of discarded states at iterations $m' > m$ (or $m > m'$),

$$\begin{aligned}
C^{(ii)}_{A(t)B} &= \sum_{l'e'm'>m} \sum_{lem} \langle lem|e^{-iHt} A e^{iHt}|l'e'm'\rangle\langle l'e'm'|B\rho|lem\rangle \\
&= \sum_{lem} \sum_{ke'} \langle lem|e^{-iHt} A e^{iHt}|ke'm\rangle\langle ke'm|B\rho|lem\rangle \\
&\approx \sum_{lm} \sum_k e^{i(E^m_k - E^m_l)t} A^m_{lk} \sum_e (B\rho)^m_{ke,le} \\
&= \sum_{lm} \sum_k e^{i(E^m_k - E^m_l)t} A^m_{lk} B^m_{kl} e^{-\beta E^m_l} \frac{w_m}{Z_m} \qquad (69) \\
C^{(iii)}_{A(t)B} &= \sum_{lem>m'} \sum_{l'e'm'} \langle lem|e^{-iHt} A e^{iHt}|l'e'm'\rangle\langle l'e'm'|B\rho|lem\rangle \\
&= \sum_{l'e'm'} \sum_{ke'} \langle l'e'm'|B\rho|ke'm\rangle\langle ke'm|e^{-iHt} A e^{iHt}|l'e'm'\rangle \\
&\approx \sum_{l'm'} (B\rho)_{l'e',ke'} e^{i(E^{m'}_{l'} - E^{m'}_k)t} A^{m'}_{kl'} \qquad (70)
\end{aligned}$$

where the NRG approximation has been used together with $\sum_e (B\rho)^m_{ke,le} = B^m_{kl} e^{-\beta E^m_l} \frac{w_m}{Z_m}$. It is also easy to show that [51],

$$\mathrm{Tr}_e \left[(B\rho)^m_{le,ke}\right] \equiv \sum_e \left[(B\rho)^m_{le,ke}\right] = \sum_{k'} B^m_{lk'} \underbrace{\sum_e \langle k'em|\rho|kem\rangle}_{R^m_{\mathrm{red}}(k',k)} \tag{71}$$

where $R^m_{\mathrm{red}}(k',k)$ is the reduced density matrix obtained from the FDM ρ by tracing out the degrees of freedom $e = (\alpha_{m+1}, ..., \alpha_N)$ [10, 51], hence the contribution $C^{(iii)}_{A(t)B}$ may be written as

$$C^{(iii)}_{A(t)B} = \sum_{lm} (B\rho)_{le,ke} e^{i(E^m_l - E^m_k)t} A^m_{kl} = \sum_{lkk'm} e^{i(E^m_l - E^m_k)t} A^m_{kl} B^m_{lk'} R^m_{\mathrm{red}}(k',k), \tag{72}$$

and

$$\begin{aligned}
C_{A(t)B} &= C^{(i)}_{A(t)B} + C^{(ii)}_{A(t)B} + C^{(iii)}_{A(t)B} \\
&= \sum_m \frac{w_m}{Z_m} \sum_{ll'} e^{i(E^m_{l'} - E^m_l)t} A^m_{ll'} B^m_{l'l} e^{-\beta E^m_l} + \sum_m \frac{w_m}{Z_m} \sum_{lk} e^{i(E^m_k - E^m_l)t} A^m_{lk} B^m_{kl} e^{-\beta E^m_l} \\
&+ \sum_m \sum_{lkk'} e^{i(E^m_l - E^m_k)t} A^m_{kl} B^m_{lk'} R^m_{\mathrm{red}}(k',k) \qquad (73)
\end{aligned}$$

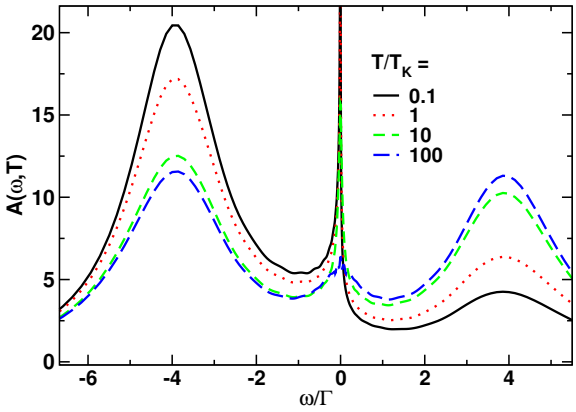


Fig. 13: *Spectral function of a negative-U quantum dot modeled by the Anderson model with $U/\Gamma = -8$ at finite gate voltage $V_g = T_{\rm K}$ and several temperatures, calculated within the FDM approach (adapted from [50]). The spectral function, polarized at low temperatures, becomes unpolarized at $T \gg T_{\rm K}$.*

Fourier transforming $-i\theta(t)C_{A(t)B}$ and the similar contribution $-i\theta(t)C_{BA(t)}$ (excercise) and adding, finally yields the Green function as

$$
\begin{aligned}
G_{AB}(\omega+i\delta) &= \sum_m \frac{w_m}{Z_m} \sum_{ll'} A^m_{ll'} B^m_{l'l} \frac{e^{-\beta E^m_l} + e^{-\beta E^m_{l'}}}{\omega + i\delta - (E^m_{l'} - E^m_l)} \\
&+ \sum_m \frac{w_m}{Z_m} \sum_{lk} A^m_{lk} B^m_{kl} \frac{e^{-\beta E^m_l}}{\omega + i\delta - (E^m_k - E^m_l)} + \sum_m \frac{w_m}{Z_m} \sum_{lk} A^m_{kl} B^m_{lk} \frac{e^{-\beta E^m_l}}{\omega + i\delta - (E^m_l - E^m_k)} \\
&+ \sum_m \sum_{lkk'} A^m_{kl} B^m_{lk'} \frac{R^m_{\rm red}(k',k)}{\omega + i\delta - (E^m_l - E^m_k)} + \sum_m \sum_{lkk'} A^m_{lk} B^m_{k'l} \frac{R^m_{\rm red}(k,k')}{\omega + i\delta - (E^m_k - E^m_l)}
\end{aligned}
\tag{74}
$$

The reduced density matrices appearing in Eq. (74) can be evaluated efficiently at all temperatures in a recursive manner [10]. The use of the complete basis set to calculate finite temperature Green functions ensures that the spectral sume rule $\int d\omega A_\sigma(\omega, T) = 1$ holds as an identity [10]. Furthermore, calculations at $\omega < T$ may be carried out without the need to restrict to a smaller cluster $M < N$, as was the case with the approach described in Sec. 4. Fig. 13 shows the spectral function of the negative-U Anderson model calculated from Eq. (74) at several temperatures.

For an application of this approach to thermoelectric properties of quantum dots see Ref. [51] and for a recent application to the magnetoresistivity and dephasing rate of multi-channel Kondo models see Ref. [31].

6 Recent Developments

The NRG has proven to be a reliable method for dealing with equilibrium properties of strongly correlated quantum impurity systems. Nevertheless, the method is still under development. In this section, we describe two areas where significant progress has been made, but where further work is needed. The first is in the transient response of a quantum impurity following either either a quantum quench, a pulse of finite duration or a periodic train of pulses [53, 33, 55, 57]. This is relevant, for example, in many pump-probe experiments [58, 59]. The second area is in developing ways to deal with real quantum impurities in metals or on surfaces, such as impurities with partially filled d- or f-levels, in which multiple channels (or bands) of the host may couple to the impurity.

Time-dependent NRG (td-NRG)

We are interested in the dynamics of a local observable $\hat{O}$ following a quantum quench in which one or more system parameters of H change suddenly at $t = 0$. Thus, the time-dependence of H is described by $H(t) = \theta(-t)H^i + \theta(t)H^f$, with H^i and H^f being time-independent initial ($t < 0$) and final state ($t > 0$) Hamiltonians, respectively [53]. The time evolution of $\hat{O}$ at $t > 0$ is then given by $O(t) = Tr\left[\rho(t)\hat{O}\right]$ where $\rho(t) = e^{iH^f t}\rho\, e^{-iH^f t}$ is the time-evolved density matrix and $\rho = e^{-\beta H^i}/Tr[\rho]$ is the density matrix of the initial state at inverse temperature β. In terms of the complete basis set, we have

$$O(t) = Tr\left[e^{iH^f t}\rho e^{-iH^f t}\hat{O}\right] = \sum_{m=m_0}^{N}\sum_{le} {}_f\langle lem|e^{-iH^f t}\rho e^{iH^f t}\hat{O}|lem\rangle_f$$

$$= \sum_{mm'=m_0}^{N}\sum_{lel'e'} {}_f\langle lem|e^{-iH^f t}\rho e^{iH^f t}|l'e'm'\rangle_{ff}\langle l'e'm'|\hat{O}|lem\rangle_f. \tag{75}$$

Making use of $1_m^+ = \sum_{m'=m+1}^{N}\sum_{l'e'}|l'e'm'\rangle\langle l'e'm'| = \sum_{ke}|kem\rangle\langle\, kem|$ [Eq. (58)], allows us to write[33]

$$\begin{aligned} O(t) &= \sum_{m=m_0}^{N}\sum_{rs\notin KK'}\sum_{e} {}_f\langle sem|e^{-iH^f t}\rho e^{iH^f t}|rem\rangle_{ff}\langle rem|\hat{O}|sem\rangle_f \\ &\approx \sum_{m=m_0}^{N}\sum_{rs\notin KK'}\sum_{e} {}_f\langle sem|e^{-iH^f_m t}\rho e^{iH^f_m t}|rem\rangle_{ff}\langle rem|\hat{O}|sem\rangle_f \\ &= \sum_{m=m_0}^{N}\sum_{rs\notin KK'}\Big(\sum_{e} {}_f\langle sem|\rho|rem\rangle_f\Big)e^{-i(E^m_s-E^m_r)t}O^m_{rs} \\ &= \sum_{m=m_0}^{N}\sum_{rs\notin KK'}\rho^{i\to f}_{sr}(m)e^{-i(E^m_s-E^m_r)t}O^m_{rs}, \end{aligned} \tag{76}$$

in which r and s may not both be kept states, $O^m_{rs} = {}_f\langle lem|\hat{O}|rem\rangle_f$ are the final state matrix elements of $\hat{O}$, which are independent of e, the NRG approximation

$$H^f|rem\rangle \approx H^f_m|rem\rangle = E^m_r|rem\rangle, \tag{77}$$

is adopted [in the second line of Eq. (76)], and $\rho^{i\to f}_{sr}(m) = \sum_e {}_f\langle sem|\rho|rem\rangle_f$ represents the reduced density matrix of the initial state projected onto the basis of final states (henceforth called the *projected density matrix*). The latter has been evaluated for the special choice of a density matrix defined on the longest Wilson chain

$$\rho = \sum_l |lN\rangle_i \frac{e^{-\beta E^N_l}}{Z_N}{}_i\langle lN|, \tag{78}$$

with $Z_N = \sum_l e^{-\beta E^N_l}$, in which only the discarded states of the last NRG iteration enter [33, 55]. More recently, the projected density matrix has been evaluated for a general initial density matrix, given by the full density matrix of the initial state [57]. This allows calculations to

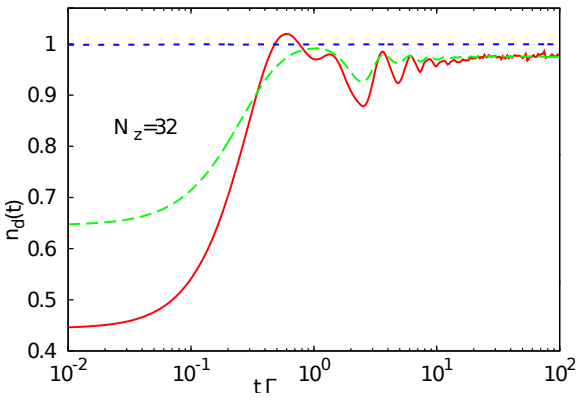


Fig. 14: *Time dependence of the occupation number $n_d(t)$ of the Anderson model for $U/\Gamma = 12$ upon switching from the mixed valence regime ($\varepsilon_d = 0$) at $t < 0$ to the symmetric Kondo regime ($\varepsilon_d = -U/2$) at $t \geq 0$ at $T = T_K/10$ (red solid line), $T = 100T_K$ (green dashed line) and $T = 10^5 T_K$ (blue dashed line). $\Lambda = 2$ with z-averaging ($N_z = 32$). Adapted from Ref. [57].*

be carried out at arbitrary finite temperature, as shown in Fig. 14. While the short-time limit $O(t \to 0^+)$ in td-NRG recovers the exact thermodynamic value $O^i = Tr[\rho O]$ in the initial state, the long-time limit suffers from an error of a few %. In addition, significant noise is observed at intermediate $t\Gamma \gtrsim 1$ to long times $t\Gamma \gg 1$. Attempts to further improve the method may be found in Ref. [57] and references therein.

Multi-orbital and Multi-channel Models for Realistic Modeling

The Anderson impurity model is a starting point for describing many different systems, from the classic examples of transition metal magnetic impurities such as Fe or Mn in non-magnetic metals such as Au, rare-earth magnetic impurities in non-magnetic metals, such as Ce in $LaAl_2$ [24], or, magnetic ions such as Co, Fe and Ti adsorbed on surfaces of non-magnetic metals such as Cu or Cu_2N/Cu (where the Cu_2N monolayer reduces the hybridization V to the substrate [60]). Of course, the relevant correlated orbitals in these systems are not the non-degenerate "s-levels" as in (3), but would be the 5-fold or 7-fold degenerate partially filled d- or f-orbitals in the case of transition metal or rare earth metal impurities, respectively. Furthermore, electrons in these partially filled shells would be subject to Coulomb, Hunds exchange, spin-orbit and crystal-field interactions, often leading to non-degenerate low energy multiplets. In addition, these low energy multiplets, would hybridize with conduction channels of appropriate symmetry, and in general with many channels, not just one as in (3). Such a non-degenerate multi-channel Anderson model capable of describing a real-transition metal impurity would then look more complicated than Eq. (3), e.g., the following model (but still neglecting spin-orbit and crystal field interactions),

$$H = \sum_{m\sigma} \varepsilon_{dm} n_{m\sigma} + \frac{1}{2} U \sum_{m\sigma} n_{m\sigma} n_{m-\sigma} + \frac{1}{2} U' \sum_{m\neq m'\sigma} n_{m\sigma} n_{m'-\sigma} + \frac{1}{2}(U'-J) \sum_{m\neq m'\sigma} n_{m\sigma} n_{m'\sigma}$$
$$- \frac{J}{2} \sum_{m\neq m'\sigma} d^\dagger_{m\sigma} d_{m-\sigma} d^\dagger_{m'-\sigma} d_{m'\sigma} - \frac{J'}{2} \sum_{m\neq m'\sigma} d^\dagger_{m\sigma} d^\dagger_{m-\sigma} d_{m'-\sigma} d_{m'\sigma}$$
$$+ \sum_{km\sigma} \epsilon_{km\sigma} c^\dagger_{km\sigma} c_{km\sigma} + \sum_{km\sigma} V_{km\sigma} (c^\dagger_{km\sigma} d_{m\sigma} + h.c.)$$

would be closer describing a real transition metal impurity such as Mn in Cu. Despite its apparent complexity, this model, just like its simpler counterpart in Eq. (3), has the same

general structure as Eq. (2) describing a general quantum impurity model, namely all many-body interactions (U, U', J, J') are contained in a local part $H_{\rm imp}$, while the multi-channel bath $H_{\rm bath}$ represents non-interacting electrons coupling via a one-body hybridization to $H_{\rm imp}$. While the NRG can be applied to such multi-channel models, for N_c-channels the Hilbert space grows as 4^{N_c} instead of 4 as for a single channel. The fraction of states that can be retained at each iteration is correspondingly smaller $(1/4^{N_c})$ than for a single channel $(1/4)$, making accurate calculations difficult, particularly for dynamical quantities. While implementing all available symmetries $(U(1), SU(2), SU(3)$, parity etc), in order to increase the fraction of states that can be retained at each iteration, will help, such symmetries are not always present. At present, reliable NRG calculations for dynamics can be carried out for 3 channel models [31]. Increasing this to 5-channels would be a significant development, allowing many interesting realistic systems to be investigated with NRG in combination with ab-initio methods to extract the relevant model parameters [52, 61, 62]. We mention here one recent proposal for achieving this, which, however, has so far only been benchmarked on a 3-channel model [63]. As in the single-band case, we rewrite the above model in linear chain form with $H_{\rm bath} = \sum_{m=1}^{N_c} \sum_{k\sigma} \epsilon_{km\sigma} c^\dagger_{km\sigma} c_{km\sigma} \rightarrow \sum_{m=1}^{N_c} \sum_{n=0}^{\infty} \sum_\sigma t_{mn} (f^\dagger_{mn\sigma} f_{mn+1\sigma} + h.c.)$, where $t_{mn} \sim D_m \Lambda^{-n/2}, m = 1, ..., N_c$ for N_c channels with half-bandwidths D_m. If all channels have the same band-width $D_m = D, m = 1, ..., N_c$, the hoppings within a shell $t_{mn}, m = 1, ..., N_c$ are constant one has to add all orbitals $f_{mn+1\sigma}, m = 1, ..., N_c$ of the next shell $n + 1$ in going from H_n to H_{n+1} in the NRG procedure of Sec. 3, before truncating the spectrum of H_{n+1}, hence leading to the above growth of the Hilbert space at each iteration. Choosing band-widths D_m with $D_1 > D_2 > ... > D_{N_c}$, as suggested in Ref. [63], leads to an energy scale separation of the orbitals within each shell, i.e., $t_{mn} \sim D_m \Lambda^{-n/2}$ for fixed n decrease with $m = 1, ..., N_c$. This allows adding the orbitals $f_{mn+1\sigma}, m = 1, ..., N_c$ of a given shell sequentially while simultaneously truncating the spectrum after each orbital is added. The calculation then resembles a single-channel calculation. The above energy scale separation is guaranteed provided $D_m/D_{m+1} = 1/g = \Lambda^{-1/2N_c}$ implying $t_{m+1n}/t_{mn} = g < 1$. Since the hoppings in this approach decrease by a factor $\Lambda^{-1/2N_c}$, a larger Λ will be required to obtain the same accuracy as a single channel calculation. In this way, the authors obtained accurate results for 3-channel and 3-impurity models.

7 Summary

Wilson's non-perturbative NRG transformation for the Kondo model has become a powerful tool for the study of quantum impurity models in general. It gives information on the many-body eigenvalues and eigenstates of such models on all energy scales and thereby allows the direct calculation of their thermodynamic, dynamic, and transport properties. Recently, it has been further developed to yield the transient response of these systems to a sudden perturbation (a quantum quench) [33], the time-dependent NRG (td-NRG). Extensions of the td-NRG to general pulses using multiple quenches have also been made [57]. The NRG also has potential to give information on the non-equilibrium steady-state transport through correlated impurity systems such as quantum dots. Recent work tries to construct a non-equilibrium density matrix for such systems by using td-NRG to time-evolve from a known initial density matrix [64].
The method has been extended in new directions, such as to models with bosonic baths to study spin-boson models [28] and the interplay of correlations and phonon effects in Anderson-Holstein models [65]. It has also been used successfully to make progress on understanding

the Mott transition, heavy fermion behaviour and other phenomena in correlated lattice systems [39, 66, 67, 68]. There is room for further improvement and extensions of the method both technically and in the investigation of more complex systems such as multi-channel models [69, 31, 63].

Acknowledgement

Useful discussions with R. Bulla, H. Nghiem and L. Merker are acknowledged (as is use of figures from joint work). Many relevant papers have not been cited in these lecture notes and we apologize to authors for this. For a more comprehensive survey of the literature on NRG work, see Ref. [4].

Appendices

A Logarithmic discretization approximation

The approximation

$$H_c = \int_{-1}^{+1} d\varepsilon\ \varepsilon\ c^+_{\varepsilon\sigma} c_{\varepsilon\sigma} \approx \sum_{n=0}^{\infty} (\varepsilon_{-n} c^+_{-n\sigma} c_{-n\sigma} + \varepsilon_{+n} c^+_{+n\sigma} c_{+n\sigma}) \tag{79}$$

used to replace the continuum band by the discrete one can be analyzed by introducing a complete orthonormal basis set of states for the conduction electrons in each interval $\pm[\Lambda^{-(n+1)}, \Lambda^{-n}]$ using the following wavefunctions

$$\psi^{\pm}_{np}(\varepsilon) = \begin{cases} \frac{\Lambda^{n/2}}{(1-\Lambda^{-1})^{1/2}} e^{\pm\omega_n p\varepsilon} & \text{for } \Lambda^{-(n+1)} < \pm\varepsilon < \Lambda^{-n} \\ 0 & \text{otherwise} \end{cases} \tag{80}$$

Here p is a Fourier harmonic index and $\omega_n = 2\pi\Lambda^n/(1-\Lambda^{-1})$. The operators $c_{\varepsilon\sigma}$ can be expanded in terms of a complete set of new operators $a_{np\sigma}, b_{np\sigma}$ labeled by the interval n and the harmonic index p

$$c_{\varepsilon\sigma} = \sum_{np} [a_{np\sigma}\psi^+_{np}(\varepsilon) + b_{np\sigma}\psi^-_{np}(\varepsilon)]. \tag{81}$$

In terms of these operators, the Kondo Hamiltonian becomes,

$$\begin{aligned} H_{KM} &= \frac{1}{2}(1+\Lambda^{-1}) \sum_{np} \Lambda^{-n} (a^\dagger_{np\sigma} a_{np\sigma} - b^\dagger_{np\sigma} b_{np\sigma}) \\ &+ \frac{(1-\Lambda^{-1})}{2\pi i} \sum_n \sum_{p\neq p'} \Lambda^{-n} (a^\dagger_{np\sigma} a_{np'\sigma} - b^\dagger_{np\sigma} b_{np'\sigma}) e^{\frac{2\pi i(p-p')}{1-\Lambda^{-1}}} \\ &+ J \sum_{\sigma\sigma'} f^+_{0\sigma} \vec{\sigma}_{\sigma\sigma'} f_{0\sigma'} . \vec{S} \end{aligned} \tag{82}$$

where in terms of the new operators, $f_{0\sigma} = \frac{1}{\sqrt{2}} \int_{-1}^{+1} d\varepsilon\ c_{\varepsilon\sigma}$ contains only $p = 0$ states:

$$f_{0\sigma} = \frac{1}{\sqrt{2}} \int_{-1}^{+1} d\varepsilon\ c_{\varepsilon\sigma} = \left[\frac{1}{2}(1-\Lambda^{-1})\right]^{1/2} \sum_{n=0}^{\infty} \Lambda^{-n/2} (a_{n0\sigma} + b_{n0\sigma}) \tag{83}$$

We notice that only the $p = 0$ harmonic appears in the local Wannier state. This is a consequence of the assumption that the Kondo exchange is independent of k. Hence the conduction electron orbitals a_{np}, b_{np} for $p \neq 0$ only couple to the impurity spin indirectly via their coupling to the a_{n0}, b_{n0} in the second term of Eq. (82). This coupling is weak, being proportional to $(1 - \Lambda^{-1})$, and vanishes in the continuum limit $\Lambda \longrightarrow 1$, so these states may be expected to contribute little to the impurity properties compared to the $p = 0$ states. This is indeed the case as shown by explicit calculations in [1, 2]. The logarithmic discretization approximation consists of neglecting conduction electron states with $p \neq 0$, resulting in H_c given by Eq. (79) with $c_{+n\sigma} \equiv a_{n,0\sigma}$ and $c_{-n\sigma} \equiv b_{n,0\sigma}$ and a discrete Kondo Hamiltonian given by Eq. (10).

B Lanczos procedure

Neglecting spin indices, the conduction electron operator is

$$H_c = \sum_k \varepsilon_k c_k^+ c_k$$

The Lanczos algorithm for tridiagonalizing this operator by repeated action on the state $|0\rangle$ is

$$|1\rangle = \frac{1}{t_0}[H_c|0\rangle - |0\rangle\langle 0|H_c|0\rangle] \tag{84}$$

$$|n+1\rangle = \frac{1}{t_n}[H_c|n\rangle - |n\rangle\langle n|H_c|n\rangle - |n-1\rangle\langle n-1|H_c|n\rangle] \tag{85}$$

yielding

$$H_c = \sum_{n=0}^{\infty} \epsilon_n f_n^+ f_n + t_n(f_n^+ f_{n+1} + H.c.) \tag{86}$$

where the site energies are given by $\epsilon_n = \langle n|H_c|n\rangle$ and the hoppings t_n are obtained as normalizations from Eqs. (84)-(85).

C Comparison with Real Space RG

Real space RG methods have been used very successfully to investigate second order phase transitions [70]. In these methods, the form of the effective Hamiltonians, H_m, is such that only a small number of couplings (e.g. nearest-neighbour and next-nearest-neighbour couplings in the case of the 2D Ising model) is retained during the RG procedure. Despite this, highly accurate results can be obtained for critical properties. The reason for this is that second order critical points are governed by just a few relevant couplings, so an effective Hamiltonian retaining just these couplings is sufficient to describe the critical behaviour. In contrast, for the Kondo model, and, for quantum impurity models in general, the interest is in obtaining information about the many-body eigenstates and eigenvalues on all energy scales and not just close to a particular fixed point where simplifying assumptions about the effective Hamiltonian might hold. Consequently, a general form of the effective Hamiltonians, including relevant and irrelevant couplings, is required in order to follow the behaviour of the system as it flows via various unstable fixed points to the stable fixed point describing the interacting quantum mechanical groundstate. Such a general form is possible in the Kondo calculation as a result of the numerical representation of the H_m.

D Comparison with DMRG

The DMRG method [22, 23], differs from the NRG approach used in the Kondo calculation in several ways. The most important, and the reason for its success as applied to one-dimensional lattice models, is the criterion for choosing the basis states of the subsystems (the "block", H_m in the Kondo calculation) used to extend the size of the system (the "superblock", H_{m+1} in the Kondo calculation). These are chosen according to their weight in a reduced density matrix built from a few eigenstates of the larger system (in the Kondo calculation this reduced density matrix would be $\rho_m^{\rm red} = \sum_{\alpha_{m+1}} \langle \alpha_{m+1}|\rho_{m+1}|\alpha_{m+1}\rangle$ where $|\alpha_{m+1}\rangle$ are the states of the $m+1$'th site and ρ_{m+1} is the density matrix of H_{m+1}). That is, the states retained in the subsystems (similar to the lowest states retained in $\bar{H}_m$ in the Kondo calculation) are in this case not necessarily the lowest energy states, but they are the states which couple most strongly, in the sense of having large eigenvalues in the reduced density matrix describing the subsystem, to the ones of interest, the target states of the larger system (in the Kondo calculation these might be taken to be the lowest few eigenstates of $\bar{H}_{m+1}$). The procedure gives highly accurate results for these target states, and therefore improves on real space NRG methods.

References

[1] Kenneth G. Wilson. The renormalization group: Critical phenomena and the kondo problem. *Rev. Mod. Phys.*, 47:773–840, Oct 1975.

[2] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson. Renormalization-group approach to the anderson model of dilute magnetic alloys. i. static properties for the symmetric case. *Phys. Rev. B*, 21:1003–1043, Feb 1980.

[3] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson. Renormalization-group approach to the anderson model of dilute magnetic alloys. ii. static properties for the asymmetric case. *Phys. Rev. B*, 21:1044–1083, Feb 1980.

[4] Ralf Bulla, Theo A. Costi, and Thomas Pruschke. Numerical renormalization group method for quantum impurity systems. *Rev. Mod. Phys.*, 80:395–450, Apr 2008.

[5] H. O. Frota and L. N. Oliveira. Photoemission spectroscopy for the spin-degenerate anderson model. *Phys. Rev. B*, 33:7871–7874, Jun 1986.

[6] T. A. Costi and A. C. Hewson. Resistivity cross-over for the non-degenerate anderson model. *Philosophical Magazine Part B*, 65(6):1165–1170, 1992.

[7] T. A. Costi, A. C. Hewson, and V. Zlatić. Transport coefficients of the anderson model via the numerical renormalization group. *J. Phys.: Condens. Matter*, 6:2519, 1994.

[8] Walter Hofstetter. Generalized numerical renormalization group for dynamical quantities. *Phys. Rev. Lett.*, 85:1508–1511, Aug 2000.

[9] Robert Peters, Thomas Pruschke, and Frithjof B. Anders. Numerical renormalization group approach to green's functions for quantum impurity models. *Phys. Rev. B*, 74:245114, Dec 2006.

[10] Andreas Weichselbaum and Jan von Delft. Sum-rule conserving spectral functions from the numerical renormalization group. *Phys. Rev. Lett.*, 99:076402, Aug 2007.

[11] D. Goldhaber-Gordon, J. Göres, M. A. Kastner, Hadas Shtrikman, D. Mahalu, and U. Meirav. From the kondo regime to the mixed-valence regime in a single-electron transistor. *Phys. Rev. Lett.*, 81:5225–5228, Dec 1998.

[12] T. A. Costi. Kondo effect in a magnetic field and the magnetoresistivity of kondo alloys. *Phys. Rev. Lett.*, 85:1504–1507, Aug 2000.

[13] B. A. Jones, C. M. Varma, and J. W. Wilkins. Low-temperature properties of the two-impurity kondo hamiltonian. *Phys. Rev. Lett.*, 61:125–128, Jul 1988.

[14] A. J. Leggett, S. Chakravarty, A. T. Dorsey, Matthew P. A. Fisher, Anupam Garg, and W. Zwerger. Dynamics of the dissipative two-state system. *Rev. Mod. Phys.*, 59:1–85, Jan 1987.

[15] U. Weiss. *Quantum dissipative systems*, volume 13. World Scientific Pub Co Inc, 2008.

[16] Antoine Georges, Gabriel Kotliar, Werner Krauth, and Marcelo J. Rozenberg. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.*, 68:13–125, Jan 1996.

[17] W.J. de Haas, J. de Boer, and G.J. van den Berg. The electrical resistance of gold, copper and lead at low temperatures. *Physica*, 1(712):1115 – 1124, 1934.

[18] P. W. Anderson. Localized magnetic states in metals. *Phys. Rev.*, 124:41–53, Oct 1961.

[19] P. W. Anderson. Poor mans derivation of scaling laws for the kondo problem. *Journal of Physics C: Solid State Physics*, 3(12):2436– 1970.

[20] R Bulla, Th Pruschke, and A C Hewson. Anderson impurity in pseudo-gap fermi systems. *Journal of Physics: Condensed Matter*, 9(47):10463, 1997.

[21] D. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner. Kondo effect in a single-electron transistor. *Nature*, 391(6663):156–159, Jan 8 1998.

[22] Steven R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, Nov 1992.

[23] U. Schollwöck. The density-matrix renormalization group. *Rev. Mod. Phys.*, 77:259–315, Apr 2005.

[24] A. C. Hewson. *The Kondo Problem to Heavy Fermions*. Cambridge University Press, Cambridge, 1997.

[25] F. Guinea, V. Hakim, and A. Muramatsu. Bosonization of a two-level system with dissipation. *Phys. Rev. B*, 32:4410–4418, Oct 1985.

[26] T. A. Costi and C. Kieffer. Equilibrium dynamics of the dissipative two-state system. *Phys. Rev. Lett.*, 76:1683–1686, Mar 1996.

[27] T. A. Costi and G. Zarand. Thermodynamics of the dissipative two-state system: A bethe-ansatz study. *Physical Review B*, 59(19):12398, 1999.

[28] Ralf Bulla, Hyun-Jung Lee, Ning-Hua Tong, and Matthias Vojta. Numerical renormalization group for quantum impurities in a bosonic bath. *Phys. Rev. B*, 71:045122, Jan 2005.

[29] F. Verstraete, V. Murg, and J. I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57:143–224, March 2008.

[30] Andreas Weichselbaum. Non-abelian symmetries in tensor networks: A quantum symmetry space approach. *Annals of Physics*, 327(12):2972 – 3047, 2012.

[31] M. Hanl, A. Weichselbaum, T. A. Costi, F. Mallet, L. Saminadayar, C. Bäuerle, and J. von Delft. Iron impurities in gold and silver: Comparison of transport measurements to numerical renormalization group calculations exploiting non-abelian symmetries. *Phys. Rev. B*, 88:075146, Aug 2013.

[32] László Borda. Kondo screening cloud in a one-dimensional wire: Numerical renormalization group study. *Phys. Rev. B*, 75:041307, Jan 2007.

[33] Frithjof B. Anders and Avraham Schiller. Real-time dynamics in quantum-impurity systems: A time-dependent numerical renormalization-group approach. *Phys. Rev. Lett.*, 95:196801, Oct 2005.

[34] L. Merker and T. A. Costi. Numerical renormalization group calculation of impurity internal energy and specific heat of quantum impurity models. *Phys. Rev. B*, 86:075150, Aug 2012.

[35] L. Merker, A. Weichselbaum, and T. A. Costi. Full density-matrix numerical renormalization group calculation of impurity susceptibility and specific heat of the anderson impurity model. *Phys. Rev. B*, 86:075153, Aug 2012.

[36] Wanda C. Oliveira and Luiz N. Oliveira. Generalized numerical renormalization-group method to calculate the thermodynamical properties of impurities in metals. *Phys. Rev. B*, 49:11986–11994, May 1994.

[37] Vivaldo L. Campo and Luiz N. Oliveira. Alternative discretization in the numerical renormalization-group method. *Phys. Rev. B*, 72:104432, Sep 2005.

[38] Osamu Sakai, Yukihiro Shimizu, and Tadao Kasuya. Single-particle and magnetic excitation spectra of degenerate anderson model with finite f–f coulomb interaction. *Journal of the Physical Society of Japan*, 58(10):3666–3678, 1989.

[39] R. Bulla, T. A. Costi, and D. Vollhardt. Finite-temperature numerical renormalization group study of the mott transition. *Phys. Rev. B*, 64:045103, Jun 2001.

[40] Axel Freyn and Serge Florens. Optimal broadening of finite energy spectra in the numerical renormalization group: Application to dissipative dynamics in two-level systems. *Phys. Rev. B*, 79:121102, Mar 2009.

[41] Rok Zitko. Adaptive logarithmic discretization for numerical renormalization group methods. *Computer Physics Communications*, 180(8):1271 – 1276, 2009.

[42] Rok Zitko and Thomas Pruschke. Energy resolution and discretization artifacts in the numerical renormalization group. *Phys. Rev. B*, 79:085106, Feb 2009.

[43] Walter Metzner and Dieter Vollhardt. Correlated lattice fermions in $d = \infty$ dimensions. *Phys. Rev. Lett.*, 62:324–327, Jan 1989.

[44] G Kotliar and D Vollhardt. Strongly correlated materials: Insights from dynamical mean-field theory. *Physics Today*, 57(3):53–59, MAR 2004.

[45] Dieter Vollhardt. Dynamical mean-field theory for correlated electrons. *Annalen der Physik*, 524(1):1, JAN 2012.

[46] M. Yoshida, M. A. Whitaker, and L. N. Oliveira. Renormalization-group calculation of excitation properties for impurity models. *Phys. Rev. B*, 41:9403–9414, May 1990.

[47] T. A. Costi. Quantum fluctuations and electronic transport through strongly interacting quantum dots. *eprint arXiv:cond-mat/0212651*, December 2002.

[48] M. Yoshida, A. C. Seridonio, and L. N. Oliveira. Universal zero-bias conductance for the single-electron transistor. *Phys. Rev. B*, 80:235317, Dec 2009.

[49] L. Merker, S. Kirchner, E. Muñoz, and T. A. Costi. Conductance scaling in kondo-correlated quantum dots: Role of level asymmetry and charging energy. *Phys. Rev. B*, 87:165132, Apr 2013.

[50] S. Andergassen, T. A. Costi, and V. Zlatić. Mechanism for large thermoelectric power in molecular quantum dots described by the negative-u anderson model. *Phys. Rev. B*, 84:241107, Dec 2011.

[51] T. A. Costi and V. Zlatić. Thermoelectric transport through strongly correlated quantum dots. *Phys. Rev. B*, 81:235127, Jun 2010.

[52] T. A. Costi, L. Bergqvist, A. Weichselbaum, J. von Delft, T. Micklitz, A. Rosch, P. Mavropoulos, P. H. Dederichs, F. Mallet, L. Saminadayar, and C. Bäuerle. Kondo decoherence: Finding the right spin model for iron impurities in gold and silver. *Phys. Rev. Lett.*, 102:056802, Feb 2009.

[53] T. A. Costi. Renormalization-group approach to nonequilibrium green functionsin correlated impurity systems. *Phys. Rev. B*, 55:3003–3009, Feb 1997.

[54] A. I. Tóth, C. P. Moca, Ö. Legeza, and G. Zaránd. Density matrix numerical renormalization group for non-abelian symmetries. *Phys. Rev. B*, 78:245109, Dec 2008.

[55] Frithjof B. Anders and Avraham Schiller. Spin precession and real-time dynamics in the kondo model:time-dependent numerical renormalization-group study. *Phys. Rev. B*, 74:245113, Dec 2006.

[56] Andreas Weichselbaum. Tensor networks and the numerical renormalization group. *Phys. Rev. B*, 86:245124, Dec 2012.

[57] H. Nghiem and T. A. Costi. Generalization of the Time-Dependent Numerical Renormalization Group Method to Finite Temperatures and General Pulses. *ArXiv e-prints*, December 2013.

[58] L. Perfetti, P. A. Loukakos, M. Lisowski, U. Bovensiepen, H. Berger, S. Biermann, P. S. Cornaglia, A. Georges, and M. Wolf. Time evolution of the electronic structure of $1t\text{-tas}_2$ through the insulator-metal transition. *Phys. Rev. Lett.*, 97:067402, Aug 2006.

[59] Sebastian Loth, Markus Etzkorn, Christopher P. Lutz, D. M. Eigler, and Andreas J. Heinrich. Measurement of fast electron spin relaxation times with atomic resolution. *Science*, 329(5999):1628–1630, 2010.

[60] A.F. Otte, M. Ternes, K. Von Bergmann, S. Loth, H. Brune, C.P. Lutz, C.F. Hirjibehedin, and A.J. Heinrich. The role of magnetic anisotropy in the kondo effect. *Nature Physics*, 4(11):847–850, 2008.

[61] P. Lucignano, R. Mazzarello, A. Smogunov, M. Fabrizio, and E. Tosatti. Kondo conductance in an atomic nanocontact from first principles. *Nature Materials*, 8:563–567, July 2009.

[62] Ryan Requist, Silvio Modesti, Pier Paolo Baruselli, Alexander Smogunov, Michele Fabrizio, and Erio Tosatti. Kondo conductance across the smallest spin 1/2 radical molecule. *Proceedings of the National Academy of Sciences*, 2013.

[63] A. K. Mitchell, M. R. Galpin, S. Wilson-Fletcher, D. E. Logan, and R. Bulla. Generalized Wilson Chain for solution of multichannel quantum impurity problems. *ArXiv e-prints*, August 2013.

[64] Frithjof B. Anders. Steady-state currents through nanodevices: A scattering-states numerical renormalization-group approach to open quantum systems. *Phys. Rev. Lett.*, 101:066804, Aug 2008.

[65] D. Meyer, A. C. Hewson, and R. Bulla. Gap formation and soft phonon mode in the holstein model. *Phys. Rev. Lett.*, 89:196401, Oct 2002.

[66] R. Bulla. Zero temperature metal-insulator transition in the infinite-dimensional hubbard model. *Phys. Rev. Lett.*, 83:136–139, Jul 1999.

[67] T. A. Costi and N. Manini. Low-energy scales and temperature-dependent photoemission of heavy fermions. *Journal of Low Temperature Physics*, 126:835–866, 2002.

[68] W. Koller, D. Meyer, and A. C. Hewson. Dynamic response functions for the holstein-hubbard model. *Phys. Rev. B*, 70:155103, Oct 2004.

[69] Th. Pruschke and R. Bulla. Hunds coupling and the metal-insulator transition in the two-band hubbard model. *The European Physical Journal B - Condensed Matter and Complex Systems*, 44(2):217–224, 2005.

[70] Kenneth G. Wilson and J. Kogut. The renormalization group and the ϵ expansion. *Physics Reports*, 12(2):75 – 199, 1974.

B 4 Introduction to Quantum Monte Carlo

Stefan Wessel
Institute for Theoretical Solid State Physics
RWTH Aachen University

Contents

1 Introduction

We just celebrated last year the 60th anniversary of the famous paper by N. Metropolis et al. [1], in which the authors outline the method of importance sampling as employed in Monte Carlo simulations of classical many-body systems. Over the years, various approaches have since then been developed that extend the importance sampling method also to models of quantum many-body systems. Various major improvements have been possible in particular for the simulation of quantum spin systems. These lecture notes aim to provide a basic introduction to the central concepts behind these modern quantum Monte Carlo (QMC) techniques for the simulation of quantum spin systems in thermal equilibrium. Acquaintance with the Monte Carlo approach to simulate classical statistical physics models is assumed, and can be obtained for example through lecture C1 by P. Mavropoulos in this volume. In devising QMC simulation methods, two major steps need to be taken: (i) The quantum partition function must be reformulated in a way as to allow stochastic sampling over a space of effective configurations, such that each contributing configuration has a positive statistical weight. (ii) An efficient update-scheme needs to be devised in order to sample this configuration space through a Markov process. The first part of these lecture notes (Secs. 2 and 3) concerns the derivation of different representations of the quantum partition function, and thus relates to the first of these steps. Similar as in the case of classical Monte Carlo schemes, a major breakthrough concerning the second step was the invention of efficient global update schemes, which will be introduced in the second part of these notes (Secs. 4 and 5). Finally, in section 6, we will discuss the QMC sign problem, which still poses the most severe restriction in applying the presented QMC methods to (frustrated) quantum spin systems. There exists already various excellent recent reviews on the methods that we discuss below, and we will mention some of these at the appropriate places. Thus, in these notes, we do not attempt to provide a comprehensive account on the subject, but instead we will highlight the key ideas behind these techniques.

2 World lines and local updates

We will discuss QMC methods for the simulation of (finite) quantum spin systems at finite temperatures. Upon tuning the simulation temperature sufficiently low (sufficiently below the finite-size spin gap), however, ground state properties of a finite system can usually be explored as well. As will be seen below, these QMC methods are feasible for non-frustrated spin models. In particular, this restricts for example the simulation of the antiferromagnetic Heisenberg model to bipartite lattices. But, there are no restrictions regarding e.g. the dimensionality of the system, and also very large systems with several thousand or even millions of spins can be treated, depending on the temperature range of interest; the computational effort of these methods scales linear in both the system size and in the inverse temperature $\beta = 1/T$ (we fix $k_B = 1$ here). For the sake of clarity, we consider in the following the spin-1/2 Heisenberg model on an open chain with N_s lattice sites,

$$\hat{H} = J_{ex} \sum_{i=1}^{N_s-1} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+1} \tag{1}$$

as an example system when deriving the QMC algorithms. We refer to appendix A concerning the notations employed here. In the following, we are interested in calculating thermal expectation values, such as for example the internal energy $E = \langle \hat{H} \rangle$ of the system at an inverse

temperature β. From statistical physics, we know that

$$\langle \hat{H} \rangle = \frac{1}{Z} \mathrm{Tr}[\hat{H} e^{-\beta \hat{H}}] = \frac{1}{Z} \sum_n E_n e^{-\beta E_n}, \tag{2}$$

where in the last step we expressed the trace in the basis of eigenstates of the Hamiltonian, and the partition function $Z = \mathrm{Tr}[e^{-\beta \hat{H}}] = \sum_n e^{-\beta E_n}$. While the above expressions in the basis of eigenstates of the Hamiltonian closely resemble the corresponding formulas for a classical system, the problem in the quantum case is that usually one does not know the full spectrum of the Hamiltonian, and thus the above sums cannot be directly calculated. Knowing the spectrum of $\hat{H}$ in the quantum case, one would often have essentially already solved the problem of interest.
QMC methods circumvent the full diagonalization of $\hat{H}$ by mapping the quantum partition function Z to a partition function of an effective classical model, and then perform a Monte Carlo sampling of the states contributing to the effective classical partition function. In order to see, how such a mapping can be realized, we consider first the Suzuki-Trotter decomposition of the partition function, as pioneered by Suzuki in 1976 [2]. This approach will set the stage for more advanced QMC approaches.

2.1 Suzuki - Trotter decomposition

Our first task will be to map the quantum partition function to that of an effective classical model. In the following, we consider in particular the quantum Heisenberg model on an open chain, given by the Hamiltonian Eq. (1), which can be decomposed into contributions from each bond,

$$\hat{H} = \sum_i \hat{H}_i, \quad \hat{H}_i = J_{ex}\, \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_{i+1} \tag{3}$$

where $\hat{H}_i$ is a bond Hamiltonian, that corresponds to the bond between site i and site $i+1$ on the chain. We can furthermore separate $\hat{H}$ into two parts,

$$\hat{H} = \hat{H}_A + \hat{H}_B, \tag{4}$$

where $\hat{H}_A$ ($\hat{H}_B$) consists of only the even (odd) bond contributions, i.e,

$$\hat{H}_A = \sum_{i\,\mathrm{even}} \hat{H}_i, \quad \hat{H}_B = \sum_{i\,\mathrm{odd}} \hat{H}_i. \tag{5}$$

While $\left[\hat{H}_A, \hat{H}_B\right] \neq 0$, the bond Hamiltonians making up $\hat{H}_A$ or $\hat{H}_B$ commute among themselves, since $[\hat{H}_i, \hat{H}_j] = 0$ for i and j both even (or both odd). We have thus separated $\hat{H}$ into two parts, each of which consists of commuting terms. Now, we rewrite the statistical operator as a product of many terms, each with a small prefactor β/M in front of $\hat{H}$,

$$e^{-\beta \hat{H}} = \left(e^{-\frac{\beta}{M}\hat{H}}\right)^M = \left(e^{-\Delta\tau \hat{H}}\right)^M, \tag{6}$$

with a (large) integer number (called the Trotter number) M, and $\Delta\tau = \beta/M$. The Suzuki-Trotter (or split operator) approximation [2, 3] now consists in approximating the exponential of H, expressed in terms of the two non-commuting pieces $\hat{H} = \hat{H}_A + \hat{H}_B$, by a product of

exponentials. While many such decompositions are possible, the most commonly used approximations (also employed beyond QMC methods) are

$$e^{-\Delta\tau\hat{H}} = e^{-\Delta\tau(\hat{H}_A+\hat{H}_B)} = \begin{cases} e^{-\Delta\tau\hat{H}_A}\cdot e^{-\Delta\tau\hat{H}_B} + O(\Delta\tau^2) \\ e^{-\Delta\tau\hat{H}_B/2}\cdot e^{-\Delta\tau\hat{H}_A}\cdot e^{-\Delta\tau\hat{H}_B/2} + O(\Delta\tau^3) \end{cases}, \tag{7}$$

where the errors are also proportional to the commutator $\left[\hat{H}_A, \hat{H}_B\right]$. Using either of the two approximations in Eq. (6), we obtain

$$\begin{aligned} Z &= \mathrm{Tr}[e^{-\beta\hat{H}}] \\ &= \mathrm{Tr}[\underbrace{e^{-\Delta\tau\hat{H}_A}\cdot e^{-\Delta\tau\hat{H}_B}}_{M}\cdot\underbrace{e^{-\Delta\tau\hat{H}_A}\cdot e^{-\Delta\tau\hat{H}_B}}_{M-1}\cdot\ldots\cdot\underbrace{e^{-\Delta\tau\hat{H}_A}\cdot e^{-\Delta\tau\hat{H}_B}}_{1}] + O(\Delta\tau^2). \end{aligned} \tag{8}$$

When using the second-order approximation from Eq. (7), the final expression is obvious. However, when using the third-order approximation, one actually obtains the same final expression, due to the cyclic invariance of the trace, which can be used to move the most left exponential to the very right, and then coalesce every other two consecutive terms. This also shows, that the systematic error in Z actually scales as $M\Delta\tau^3 \propto \Delta\tau^2$, even when using the second-order approximation, where one might have expected an $O(\Delta\tau)$ error in the final expression. Now, consider a basis of the Hilbert space, e.g. in terms of the local eigenstates of $\hat{S}_i^z$, which we write as

$$\begin{aligned} |\sigma_1\rangle &= |\uparrow\uparrow\ldots\uparrow\uparrow\rangle \\ |\sigma_2\rangle &= |\uparrow\uparrow\ldots\uparrow\downarrow\rangle \\ &\vdots \\ |\sigma_{2^N}\rangle &= |\downarrow\downarrow\ldots\downarrow\downarrow\rangle. \end{aligned} \tag{9}$$

Since this set forms a basis, we obtain a completeness relation:

$$\sum_\sigma |\sigma\rangle\langle\sigma| = 1, \tag{10}$$

where $|\sigma\rangle\langle\sigma|$ is a projection operator onto the basis state $|\sigma\rangle$. Within this basis of the Hilbert space, we can thus express the partition function as

$$\begin{aligned} Z &= \sum_{\sigma^0}\langle\sigma^0|e^{-\beta\hat{H}}|\sigma^0\rangle \\ &\approx \sum_{\sigma^0}\langle\sigma^0|\underbrace{e^{-\Delta\tau\hat{H}_A}e^{-\Delta\tau\hat{H}_B}}_{M}\underbrace{e^{-\Delta\tau\hat{H}_A}e^{-\Delta\tau\hat{H}_B}}_{M-1}\cdots\underbrace{e^{-\Delta\tau\hat{H}_A}e^{-\Delta\tau\hat{H}_B}}_{1}|\sigma^0\rangle \\ &= \sum_{\sigma^0}\langle\sigma^0|e^{-\Delta\tau\hat{H}_A}e^{-\Delta\tau\hat{H}_B}\ldots e^{-\Delta\tau\hat{H}_A}\left(\sum_{\sigma^1}|\sigma^1\rangle\langle\sigma^1|\right)e^{-\Delta\tau\hat{H}_B}|\sigma^0\rangle \\ &= \sum_{\sigma^0}\sum_{\sigma^1}\langle\sigma^0|e^{-\Delta\tau\hat{H}_A}e^{-\Delta\tau\hat{H}_B}\ldots\left(\sum_{\sigma^2}|\sigma^2\rangle\langle\sigma^2|\right)e^{-\Delta\tau\hat{H}_A}|\sigma^1\rangle\langle\sigma^1|e^{-\Delta\tau\hat{H}_B}|\sigma^0\rangle \\ &= \sum_{\sigma^0,\sigma^1,\sigma^2}\langle\sigma^0|e^{-\Delta\tau\hat{H}_A}e^{-\Delta\tau\hat{H}_B}\ldots\left(\sum_{\sigma^3}|\sigma^3\rangle\langle\sigma^3|\right)e^{-\Delta\tau\hat{H}_B}|\sigma^2\rangle\langle\sigma^2|e^{-\Delta\tau\hat{H}_A}|\sigma^1\rangle\langle\sigma^1|e^{-\Delta\tau\hat{H}_B}|\sigma^0\rangle. \end{aligned}$$

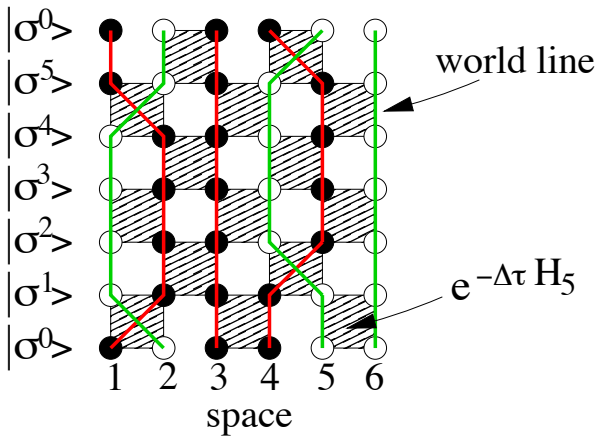


Fig. 1: *A configuration contributing to the checkerboard decomposition of the partition function of an open spin-*$1/2$ *Heisenberg chain with 6 lattice sites, for a Trotter number* $M = 3$*. Black (white) circles denote spin up (down), and the world lines of the up (down) spins are denoted by red (green) lines. Each dashed square represents the exponential of a bond Hamiltonian.*

In the above summations, we added a (superscript) label to distinguish the basis sets that arise from the trace ($|\sigma^0\rangle$) and from the various inserted partitions of unity ($|\sigma^1\rangle$, $|\sigma^2\rangle$, and $|\sigma^3\rangle$). Continuing this way, we eventually arrive at the following expression:

$$Z = \sum_{\{\sigma^i\}} \underbrace{\langle\sigma^0|e^{-\Delta\tau\hat{H}_A}|\sigma^{2M-1}\rangle}_{2M} \underbrace{\langle\sigma^{2M-1}|e^{-\Delta\tau\hat{H}_B}|\sigma^{2M-2}\rangle}_{2M-1} \underbrace{\langle\sigma^{2M-2}|e^{-\Delta\tau\hat{H}_A}|\sigma^{2M-3}\rangle}_{2M-2} \ldots$$
$$\ldots \underbrace{\langle\sigma^3|e^{-\Delta\tau\hat{H}_B}|\sigma^2\rangle}_{3} \underbrace{\langle\sigma^2|e^{-\Delta\tau\hat{H}_A}|\sigma^1\rangle}_{2} \underbrace{\langle\sigma^1|e^{-\Delta\tau\hat{H}_B}|\sigma^0\rangle}_{1} + O(\Delta\tau^2). \quad (11)$$

This representation of Z is referred to as the "Suzuki-Trotter decomposition". Furthermore, since $\hat{H}_A$ and $\hat{H}_B$ each consists of commuting parts, find that

$$e^{-\Delta\tau\hat{H}_A} = \prod_{i \text{ even}} e^{-\Delta\tau\hat{H}_i} = e^{-\Delta\tau\hat{H}_2} e^{-\Delta\tau\hat{H}_4} \ldots \quad (12)$$

$$e^{-\Delta\tau\hat{H}_B} = \prod_{i \text{ odd}} e^{-\Delta\tau\hat{H}_i} = e^{-\Delta\tau\hat{H}_1} e^{-\Delta\tau\hat{H}_3} \ldots \quad (13)$$

and thus in the Suzuki-Trotter decomposition, each exponential expression factorizes into exponentials for the bond Hamiltonians on either all even or all odd bonds. One can represent a given contribution to Z graphically, such as shown for a specific example with $N_s = 6$, and $M = 3$ in Fig. 1. This two-dimensional picture appears like a space-time picture of spins propagating in discrete steps from the initial configuration $|\sigma^0\rangle$ to $|\sigma^1\rangle$,..., and finally from $|\sigma^{2M-1}\rangle$ back to $|\sigma^0\rangle$, since the configurations on the first and the last step are equal. This pictorial illustration also explains why the underlying partitioning of the Hamiltonian is often referred to as the "checkerboard decomposition" .

2.2 World lines

Each exponential $e^{-\Delta\tau\hat{H}_i}$ couples two spin sites on a shaded plaquette, and the statistical weight of a configuration is the product of all the matrix elements of the exponentials on all shaded

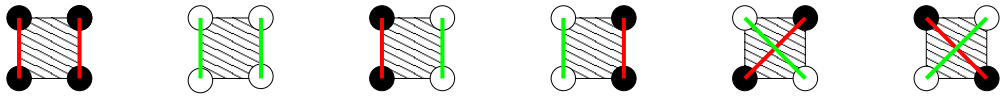


Fig. 2: *Allowed shaded plaquette configurations for the spin-1/2 Heisenberg model. Black (white) circles denote spin up (down), and the world lines of the up (down) spins are denoted by red (green) lines.*

plaquettes. Note, that $e^{-\Delta\tau \hat{H}_i}$ appears as an imaginary-time evolution operator, that propagates the two spins at sites i and $i+1$ from one "time slice" to the next. To appreciate this suggestive picture of a discrete time propagation, we introduced the corresponding "Trotter time-step" $\Delta\tau = \beta/M$ above. However, not all of the $2^4 = 16$ possible spin configurations along a shaded plaquette are allowed to appear within the checkerboard decomposition. In order to assess the allowed configurations and the corresponding matrix elements, we explicitly calculate the matrix elements of the operators

$$e^{-\Delta\tau \hat{H}_i} = \sum_{k=0}^{\infty} \frac{1}{k!} (\Delta\tau)^k (-\hat{H}_i)^k. \tag{14}$$

In the local two-sites basis of the two spins at site i and $i+1$, the Hamiltonian matrix is

$$\hat{H}_i = J_{ex} \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{pmatrix}, \quad \text{in the local basis} \quad \begin{matrix} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{matrix}. \tag{15}$$

Upon performing the Taylor expansion, and grouping back the resulting terms, one finds (this would make up a nice exercise)

$$e^{-\Delta\tau \hat{H}_i} = e^{+\Delta\tau J_{ex}/4} \begin{pmatrix} e^{\frac{-\Delta\tau J_{ex}}{2}} & 0 & 0 & 0 \\ 0 & \cosh(\frac{\Delta\tau J_{ex}}{2}) & -\sinh(\frac{\Delta\tau J_{ex}}{2}) & 0 \\ 0 & -\sinh(\frac{\Delta\tau J_{ex}}{2}) & \cosh(\frac{\Delta\tau J_{ex}}{2}) & 0 \\ 0 & 0 & 0 & e^{\frac{-\Delta\tau J_{ex}}{2}} \end{pmatrix}. \tag{16}$$

There are thus 6 possible allowed plaquettes, shown in Fig. 2, with a finite weight. All other combinations would lead to a vanishing matrix element and are thus not allowed to occur as part of any allowed configuration. The allowed configurations are those that exhibit the same total local magnetization $S_i^z + S_{i+1}^z$ on the lower and the upper edge of each shaded plaquette. This property derives from the fact that each bond Hamiltonian $\hat{H}_i$ conserves the total magnetization of the two spins connected by the bond.

When we connect the positions of the up and down spins as they propagate through the shaded plaquettes, we obtain continuous lines. These are the "world lines" of the spins, and each such world line denotes the evolution of one spin up or down from $|\sigma^0\rangle$ back to $|\sigma^0\rangle$. The world lines for both the spin-up and the spin-down case are illustrated in Fig. 1. Note, that it would be sufficient to only show, say, the spin-up spins, which is what will be done later on in these notes. If one would glue together the upper and lower boundary of the space-time configuration to make the periodicity constraint by the trace more explicit, these world lines are thus continuous,

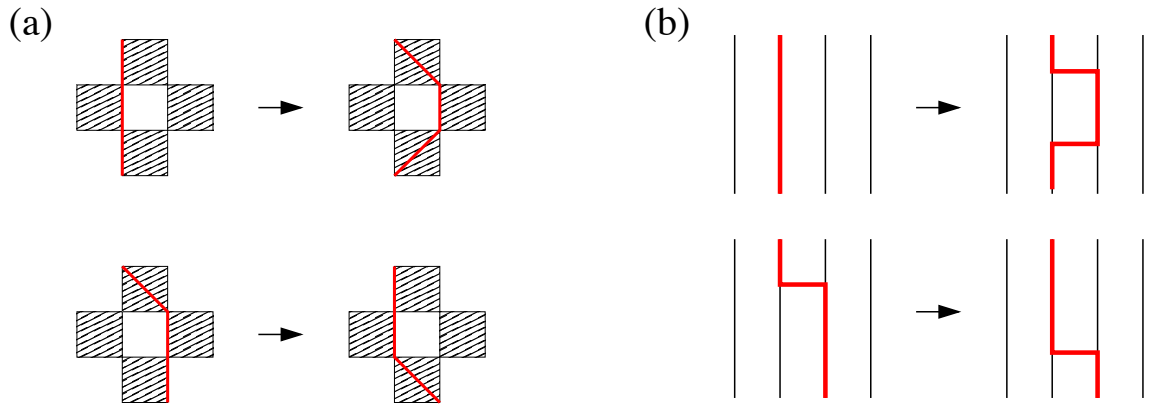


Fig. 3: *Two different local updates of world-line configurations in discrete time (a) and in the continuous-time limit (b). For clarity, only the spin-up world lines are shown explicitly.*

and are not broken anywhere. This set of unbroken world-line configurations defines a classical statistical model, which has the same partition function as the quantum partition function, if the weight of a given world-line configuration C equals the product of the matrix elements from all shaded plaquettes P:

$$W(C) = \prod_P W_P(C|_P), \tag{17}$$

where $W_P(C|_P)$ is the corresponding matrix element of $e^{-\Delta\tau \hat{H}_i}$ for plaquette P in the configuration C. At first sight, the effective model might appear to be just the two-dimensional Ising model on a square lattice of size $N_s \times 2M$, with periodic boundary conditions in the y-direction. However, the effective classical model is in fact more complex than the Ising model, which has an unconstrained configuration space. Only a subset of configurations of the two-dimensional Ising model are allowed to occur also in the effective classical model for Z. These configurations are those that correspond to continuous, unbroken world lines. In higher spatial dimensions, one can use the Suzuki-Trotter approach with decompositions very similar to the one discussed above. One thus finds that the quantum partition function of a d-dimensional system (here $d = 1$) corresponds to that of an effective classical model in $(d+1)$ dimensions, with the $(d+1)$-th direction corresponding to an imaginary-time evolution of the original quantum model. This well known quantum-to-classical-mapping holds in fact much more generally, and exhibits a deeper connection between classical and quantum statistical physics.
Note however the minus sign in front of the off-diagonal matrix elements: for the allowed world-line configurations, all the accumulated signs actually cancel out, due to the periodicity constraint in imaginary time, in case of the open chain or for a closed chain with an even number of sites. This holds also true for higher-dimensional generalizations of the world-line approach whenever the underlying lattice structure is bipartite. However, for frustrated systems, e.g. for a closed 3-site chain, there appears a "sign problem", since both configuration with positive and negative weights are allowed. We will discuss the consequences of this issue in Sec. 6.

2.3 Local updates

The world-line representation introduced above can be taken as a starting point to set up a QMC algorithm [4]. But, how does one sample configurations in the effective world-line model in a

Monte Carlo algorithm, i.e., how do we generate new valid world-line configurations from a given one? Since the spin $\hat{S}^z$-conservation prohibits the breaking of world lines, the updates need to move world-lines segment instead of just changing local spin states like in the classical Ising model. In early days, so-called local updates were performed, i.e. local manipulations of the world-line configuration, such as those shown in Fig. 3 (a). These local updates are quite simple and either generate a pair of kinks across a white plaquette within a straight section of the world line, or they move a kink to a neighboring plaquette [4, 5], as shown in Fig. 3 (a). Slightly more complicated local moves are required for higher-dimensional spin systems [6], or for other quantum lattice models [7]. In any case, such local updates are accepted by considering the related change in the statistical weight according to Eq. (17), using e.g. the Metropolis [1] acceptance scheme.

However, such local updates are not efficient in two respects: First, global properties of the configurations cannot be changed by such local updates; the spatial winding number of the world lines and their total number, i.e. the total magnetization of the system, cannot be changed using local moves. Thus, they had to be complemented by special global updates as well [6]. Second, such updates lead to severe critical slowing down upon approaching critical regions, similarly as those familiar from the local Metropolis algorithm for the Ising model. A major breakthrough in overcoming this problem was the work by Evertz, Lana and Marcu [8], in which they presented an extension of the Swendsen-Wang cluster update idea known from classical Monte Carlo studies to world-line QMC methods [56]. We will present this "loop algorithm" in Sec. 4. Before doing so, we want to introduce two other QMC representations, which are relevant for modern QMC algorithms.

2.4 The continuous-time limit

Due to the finite number of time slices, M, the approach described above suffers from a systematic error – the discretization or Trotter error. It has been shown that in most cases, one can keep $\Delta\tau$ independent of N_s and β in order to ensure a constant error level [9] (for some observables however, care has to be taken to avoid divergent errors in the zero temperature limit $\beta \to \infty$[9, 10]). The Trotter error was controlled originally by extrapolation to the continuous-time limit $\Delta\tau \to 0$, from simulations with different values of the time step $\Delta\tau$. It was realized later [11] that the continuous-time limit can be taken already in the construction of the algorithm, so that simulations can be performed directly in the limit $\Delta\tau \to 0$, i.e. $M \to \infty$, without the need to perform any final extrapolations. This appears feasible, once one realizes that a given world-line configuration can be represented by keeping only a list of times at which the configuration *changes*, instead of storing the configuration at each of the $2M$ time slices. Indeed, the probability for a jump of a world line (i.e., a kink in the world line) from one lattice site to a neighboring site across a given plaquette, and thus a change in the local configuration is proportional to $\sinh(\Delta\tau J_{ex}/2) \propto \Delta\tau \propto 1/M$. Hence, the mean total number of such kinks remains finite in the limit $M \to \infty$. The continuous world-line representation, i.e. the limit $\Delta\tau \to 0$, is thus well defined, and the relevant configurational information can be efficiently represented on a computer. This is illustrated in Fig. 4 (a) and (b), by comparing the discrete and continuous-time world-line approach.

In addition to the configurational information, the local updates need to be revisited in the continuous-time limit as well. In particular, the probability P_{pk} for the insertion of a pair of

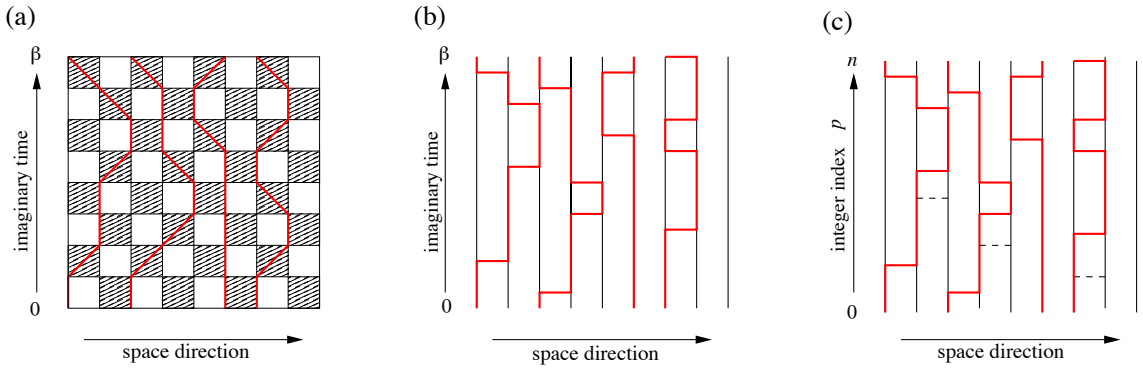


Fig. 4: *Comparison of world-line QMC configurations in a) discrete time, b) continuous time, and c) in the SSE representation. In the last case, the continuous-time index is replaced by an ordered integer index of the operators, and additional diagonal terms are indicated by dashed lines.*

kinks in the world line (the upper move in Fig. 3 (a)), vanishes in the continuous-time limit as

$$P_{pk} = \sinh^2(\Delta\tau J_{ex}/2)/\cosh^2(\Delta\tau J_{ex}/2) \propto \Delta\tau^2 \propto 1/M^2 \to 0. \tag{18}$$

To circumvent such a vanishing probability, one now proposes to insert a pair of jumps not at specific locations, but *anywhere* inside a finite time interval [11], as illustrated in Fig. 3 (b). The integrated probability for such a move then remains finite in the continuous-time limit. Similarly, instead of shifting a kink by $\Delta\tau$ (the lower move in Fig. 3 (a)), one now moves it within a finite time interval in the continuous-time algorithm. In addition to local updates, also the loop algorithm has been shown to allow for an efficient realization in the continuous-time limit [12], as will be discussed below.
Finally, we remark that the continuous-time limit of the Suzuki-Trotter formula, Eq. (11), in the above interpretation is in fact equivalent to a time-dependent perturbation theory in imaginary time,

$$\begin{aligned} Z &= \mathrm{Tr}\exp(-\beta\hat{H}) = \mathrm{Tr}\left[\exp(-\beta\hat{H}_0)\,\mathcal{T}\exp\int_0^\beta d\tau\hat{V}(\tau)\right] \\ &= \mathrm{Tr}\left[\exp(-\beta\hat{H}_0)\left(1-\int_0^\beta d\tau\hat{V}(\tau)d\tau+\frac{1}{2}\int_0^\beta d\tau_1\int_{\tau_1}^\beta d\tau_2\hat{V}(\tau_1)\hat{V}(\tau_2)+\ldots\right)\right], \end{aligned} \tag{19}$$

where the symbol $\mathcal{T}$ denotes the (imaginary-) time ordering of the exponential [11, 13]. In this representation, the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$ is split up into a diagonal term $\hat{H}_0$ and an off-diagonal perturbation $\hat{V}$, which in the interaction picture is $\hat{V}(\tau) = \exp(\tau\hat{H}_0)\hat{V}\exp(-\tau\hat{H}_0)$. For the Heisenberg model, the diagonal term $\hat{H}_0$ is given by the longitudinal $\hat{S}_i^z\hat{S}_j^z$ spin-spin interactions, while the off-diagonal perturbation $\hat{V}$ relates to the transverse, spin exchange terms, $\frac{1}{2}(\hat{S}_i^+\hat{S}_j^- + \hat{S}_i^-\hat{S}_j^+)$ in $\hat{H}$. In more detail, we express the Heisenberg Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{V} = \sum_{\langle i,j\rangle} J_{ex}\hat{S}_i^z\hat{S}_j^z + \sum_{\langle i,j\rangle}\frac{J_{ex}}{2}(\hat{S}_i^+\hat{S}_j^- + \hat{S}_i^-\hat{S}_j^+). \tag{20}$$

The number of kinks in a given continuous-time world-line configuration is then equal to the expansion order of a specific term in the perturbation expansion, wherein each kink's space-time position is set by a specific spin exchange term from $\hat{V}$ between two lattice sites i and j,

as well as a corresponding imaginary time τ between 0 and β. Furthermore, the exponential factors $\exp(\tau\hat{H}_0)$ in $\hat{V}(\tau)$ describe the vertical, unperturbed evolution of the spins along the world lines between the kinks, thus establishing explicitly the above-mentioned equivalence between these two formulations of the continuous-time limit.

3 Stochastic series expansion

Already before Suzuki's approach to QMC in the mid 1970s [2, 4], an alternative approach to QMC simulations has been put forward by Handscomb in the early 1960s [14, 15] for the specific case of the ferromagnetic Heisenberg model, which is based on a Taylor expansion of the statistical operator inside the partition function:

$$Z = \mathrm{Tr}\exp(-\beta\hat{H}) = \sum_{n=0}^{\infty}\frac{\beta^n}{n!}\mathrm{Tr}[(-\hat{H})^n]. \tag{21}$$

In Handscomb's approach and within later extensions to other models [16, 17, 18], the traces of H^n where evaluated employing projection operator expressions. Later, in the early 1990s, the power-series approach to QMC simulations was revisited by Sandvik and Kurkijärvi within the more generally applicable stochastic series expansion (SSE) formulation [19, 20], in which these traces are also sampled stochastically. A recent review of the SSE method, combined with a general introduction to computational methods for quantum spin systems, can be found in Ref. [21]. It contains also many basic details for implementing the SSE QMC method.

3.1 Configuration space

To formulate the SSE method, it proves convenient to first express the Heisenberg Hamiltonian as a sum of bond operators, which are either diagonal or off-diagonal in the standard $\hat{S}^z$-basis,

$$\hat{H} = -\sum_{b=1}^{N_b}(\hat{H}_{1,b} - \hat{H}_{2,b}) + J_{ex}N_b/4, \tag{22}$$

with

$$\hat{H}_{1,b} = J_{ex}\left(\frac{1}{4} - \hat{S}^z_{i(b)}\hat{S}^z_{j(b)}\right), \tag{23}$$

$$\hat{H}_{2,b} = \frac{J_{ex}}{2}\left(\hat{S}^+_{i(b)}\hat{S}^-_{j(b)} + \hat{S}^-_{i(b)}\hat{S}^+_{j(b)}\right). \tag{24}$$

Here, N_b equals the total number of bonds in the system, and $i(b)$ and $j(b)$ denote the two lattice sites that are connected by the bond with bond index b. The explicit introduction of the minus sign in $\hat{H}$ is convenient, while the constant in the diagonal operators will make the series expansion positive definite, as we will see shortly. Note that this constant is irrelevant for the physics of the system, but has to be accounted for when calculating the system's energy, as we will see in the next section. We now insert the above form of the Hamiltonian into the expression Eq. (21) for the partition function, thereby obtaining

$$Z = \sum_{\sigma}\sum_{n=0}^{\infty}\frac{\beta^n}{n!}\sum_{\mathcal{S}_n}(-1)^{n_2}\langle\sigma|\hat{H}_{t(n),b(n)}\cdots\hat{H}_{t(2),b(2)}\hat{H}_{t(1),b(1)}|\sigma\rangle. \tag{25}$$

Here, we used the local $\hat{S}^z$-basis to express the trace, and $\mathcal{S}_n$ denotes products (strings) of the bond operators $\hat{H}_{1,b}$ or $\hat{H}_{2,b}$ that originate from expanding $\hat{H}^n$ in these bond operator terms using Eq. (22), namely, each such operator string $\mathcal{S}_n$ is an sequence of n bond operators, specified by the type labels $t(p) \in \{1, 2\}$ (i.e. diagonal or off-diagonal) and $b(p) \in \{1, ..., N_b\}$, so that we can also write such a sequence as

$$\mathcal{S}_n = [t(1), b(1)], [t(2), b(2)], ..., [t(n), b(n)]. \tag{26}$$

The summation in Eq. (25) extends over all expansion orders n, and for each expansion order, over all operator sequences $\mathcal{S}_n$ of length n, containing n bond operators. Furthermore, n_2 denotes the number of offdiagonal operators in the sequence $\mathcal{S}_n$, i.e. the number of sequence elements with $t(p) = 2$. It is important to note, that in the local $\hat{S}^z$-basis, the bond operators exhibit a non-branching property, i.e., if one applies a bond operator to one of these basis states, the resulting state is either proportional to the same basis state, or to another basis state, or it vanishes. But in no case is a superposition of two or more such basis states generated when traversing the action of the sequence of bond operators starting from any initial basis state $|\sigma\rangle$. The propagated basis state that appears after the action of the first p operators (i.e. at propagation level p) will be denoted (after proper normalization) by $|\sigma(p)\rangle$, i.e.,

$$|\sigma(p)\rangle \propto \prod_{q=1}^{p} \hat{H}_{t(q),b(q)} |\sigma\rangle. \tag{27}$$

Note that for any operator-state configuration $(\mathcal{S}_n, |\sigma\rangle)$ that contributes to Z, the final state, resulting after the action of all the n operators has to fulfill $\langle\sigma|\sigma(n)\rangle \neq 0$, thus we obtain the following periodicity constraint: $|\sigma(n)\rangle = |\sigma(0)\rangle = |\sigma\rangle$.
If one examines in more detail the action of the different bond operators $\hat{H}_{1,b}$ and $\hat{H}_{2,b}$ on the two spins at the related lattice sites $i(b)$ and $j(b)$ of a basis state $|\sigma\rangle$, one finds, that (i) the state is destroyed if both spins are parallel,

$$\hat{H}_{1,b}|\uparrow_{i(b)}\uparrow_{j(b)}\rangle = 0, \quad \hat{H}_{1,b}|\downarrow_{i(b)}\downarrow_{j(b)}\rangle = 0, \tag{28}$$

$$\hat{H}_{2,b}|\uparrow_{i(b)}\uparrow_{j(b)}\rangle = 0, \quad \hat{H}_{2,b}|\downarrow_{i(b)}\downarrow_{j(b)}\rangle = 0, \tag{29}$$

and (ii) if the two spins are anti-parallel, the spin state is preserved by $\hat{H}_{1,b}$, while being flipped by $\hat{H}_{2,b}$, in both cases with a matrix element of $J_{ex}/2$,

$$\langle\uparrow_{i(b)}\downarrow_{j(b)}\ |\hat{H}_{1,b}|\uparrow_{i(b)}\downarrow_{i(b)}\rangle = J_{ex}/2, \quad \langle\downarrow_{i(b)}\uparrow_{j(b)}\ |\hat{H}_{1,b}|\downarrow_{i(b)}\uparrow_{i(b)}\rangle = J_{ex}/2, \tag{30}$$

$$\langle\downarrow_{i(b)}\uparrow_{j(b)}\ |\hat{H}_{2,b}|\uparrow_{i(b)}\downarrow_{i(b)}\rangle = J_{ex}/2, \quad \langle\uparrow_{i(b)}\downarrow_{j(b)}\ |\hat{H}_{2,b}|\downarrow_{i(b)}\uparrow_{i(b)}\rangle = J_{ex}/2. \tag{31}$$

From (i) it follows that for any operator-state configuration $(\mathcal{S}_n, |\sigma\rangle)$ that contributes to Z, the bond operators $\hat{H}_{t(p),b(p)}$ act only on propagated states with anti-parallel spins on bond $b(p)$. Property (ii) implies that all the matrix elements that arise from the bond operators within the operator string are equal and positive. The effective configuration space of the SSE method thus essentially consists of all allowed operator-state configurations (i.e., those with non-zero weight), with the weight of a given configuration $C = (\mathcal{S}_n, |\sigma\rangle)$ obtained from Eq. (25),

$$Z = \sum_C W(C), \quad W(C) = \frac{\beta^n}{n!} \langle\sigma| \prod_{p=1}^{n} H_{t(p),b(p)} |\sigma\rangle. \tag{32}$$

Here, we also anticipated that on a bipartite lattice the number of off-diagonal operators in an allowed operator sequence must be even, such that $(-1)^{n_2} = +1$, due to the periodicity constraint $|\sigma(n)\rangle = |\sigma(0)\rangle = |\sigma\rangle$. For a non-bipartite (frustrated) lattice this would not be the case, and a QMC sign-problem would result; cf. Sec. 6 for a discussion of this case. The SSE QMC configurations can be visualized very similar as in the previous case, cf. Fig. 4 for a comparison between the different formulations. Note that in the SSE formulation a discrete time-like index p is introduced. Furthermore, in addition to the jump events that appear in the continuous-time world-line formulation, the SSE configurations exhibit the additional presence of diagonal bond-operators. Indeed, the SSE representation can be formally related to the continuous-time world-line representation by observing that Eq. (25) is obtained from Eq. (19) upon setting $\hat{H}_0 = 0$, $\hat{V} = \hat{H}$ and integrating over all times (compare also Fig. 4) [22]. This mapping exposes the respective advantages and disadvantages of the two representations: The SSE representation corresponds to a perturbation expansion in all terms of the Hamiltonian, whereas continuous-time world-line algorithms treat the diagonal terms exactly and perturbs only in the off-diagonal terms of the Hamiltonian. The continuous-time algorithm hence needs fewer terms in the expansion, but we pay for it by having to handle continuous imaginary-time variables. For spin systems in which the contributions to the total energy arising from diagonal and off-diagonal terms are well balanced, the SSE is typically preferable in practice, due to its discrete representation of the imaginary-time continuum. It is important to note, that this discretization does *not* introduce any Trotter-error as in the discrete-time world-line QMC; instead, the SSE formulation should be considered an essentially exact QMC technique.

It is possible to formulate Monte Carlo sampling algorithms for the SSE configuration space introduced above, i.e., containing operator strings of fluctuating length [14, 20, 23]. However, it is also possible to work with operator strings of a fixed length, which is in most cases more convenient computationally [19, 20]. Indeed,it will be shown in the next section, the mean length of the operator string $\langle n\rangle$ is essentially related to the energy of the system and scales linear with the system size N_s and β. Furthermore, the specific heat is given by $C = \langle n^2\rangle - \langle n\rangle^2 - \langle n\rangle$, so that in particular for $T \to 0$, where C vanishes, the variance of the distribution of expansion orders is seen to be equal to $\langle n\rangle$. Hence, for a given system size and temperature, the series expansion order n of the operator-state configurations $C = (\mathcal{S}_n, |\sigma\rangle)$ can in practice be truncated at a sufficiently large cutoff Λ, which is typically determined within the thermalization phase of an SSE simulation, without introducing any detectable error. It is then feasible to work with fixed-length operator strings upon augmenting all operator sequences of length $n < \Lambda$ by $\Lambda - n$ unit operators $\hat{I}$, denoted in the following by $\hat{H}_{0,0} = \hat{I}$. Allowing for all possible positions of the unit operators in the original operator strings, one is then lead to the following representation of the partition function:

$$Z = \sum_C W_\Lambda(C), \quad W(C) = \frac{\beta^n(\Lambda - n)!}{\Lambda!} \langle\sigma| \prod_{p=1}^{\Lambda} \hat{H}_{t(p),b(p)} |\sigma\rangle, \tag{33}$$

where each operator-state configuration $(\mathcal{S}_\Lambda, |\sigma\rangle)$ now consists of a sequence of operators specified by the type labels $t(p) \in \{0, 1, 2\}$ and $b(p) \in \{1, ..., N_b\}$ in case $t(p) \in \{1, 2\}$, or $b(p) = 0$, if $t(p) = 0$, which would represent a unit operator. In the above formula, n denotes the number of non-unit operators in the operator sequence, i.e. the number of true bond operators from $\hat{H}$. The new combinatorial factor in $W_\Lambda(C)$ accounts for the $\binom{\Lambda}{n}$ equivalent terms that are generated this way as compared to the expression without unit operators in Eq. (25). Again, we left out the factor $(-1)^{n_2} = +1$ on bipartite lattices, where all allowed configurations thus contribute

with a positive weight $W_\Lambda(C) > 0$ to the partition function, which is hence feasible for Monte Carlo sampling.

3.2 Observables

Thus far in this exposition of QMC methods, we focused on different representations of the quantum partition function in terms of an effective set of configurations that prove amenable to Monte Carlo sampling. However, we did not mention yet how observables can be calculated. Here, we provide an overview how typical observables are measured within the SSE framework [19, 20, 24]. In many cases, the generalization from the formulas provided below to the other world-line methods is straightforward. A remarkably simple formula holds for the total inner energy $E = \langle \hat{H} \rangle$, which can be expressed in terms of the mean expansion order:

$$\langle \hat{H} \rangle = \frac{1}{Z} \mathrm{Tr}[\hat{H} e^{-\beta \hat{H}}] = \frac{1}{Z} \sum_{n=0}^{\infty} \frac{\beta^n}{n!} \mathrm{Tr}[(-\hat{H})^{n+1}] = -\frac{1}{Z} \sum_{n=1}^{\infty} \frac{\beta^n}{n!} \frac{n}{\beta} \mathrm{Tr}[(-\hat{H})^n] = -\frac{\langle n \rangle}{\beta}. \quad (34)$$

Accounting for the constant shift that we added to the Hamiltonian in Eq. (22) to perform the SSE expansion, we thus obtain the estimator for the inner energy before the shift as

$$E = -\frac{\langle n \rangle}{\beta} + J_{ex} N_b / 4. \quad (35)$$

By the same procedure, that lead to Eq. (34), one finds that $\langle \hat{H}^2 \rangle = \langle n(n-1) \rangle / \beta^2$, from which the specific heat estimator

$$C = \beta^2 (\langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2) = \langle n^2 \rangle - \langle n \rangle^2 - \langle n \rangle \quad (36)$$

is readily obtained, which we already advocated above. Note, that the constant introduced in Eq. (22) cancels out in the estimator for C, as would be expected. In practice, the statistical error of C becomes large at low temperatures (where C itself becomes small), since it is given by the difference of two large numbers ($\sim (N_s \beta)^2$). We can also readily obtain the expectation value of any operator $\hat{A}$ that is diagonal in the local $\hat{S}^z$-basis, such as for example a spin-spin correlation function $\langle \hat{S}_i^z \hat{S}_j^z \rangle$ (in which case $\hat{A} = \hat{S}_i^z \hat{S}_j^z$), by using averages over the propagated states:

$$\langle \hat{A} \rangle = \left\langle \frac{1}{n} \sum_{p=0}^{n-1} A(p) \right\rangle, \quad (37)$$

where $A(p) = \langle \sigma(p) | \hat{A} | \sigma(p) \rangle$ has been introduced for convenience. From the spin-spin correlation function, the equal-time structure factor is then obtained by Fourier transformation.
Finally, let us mention, how one can measure generalized susceptibilities related to two diagonal operators $\hat{A}, \hat{B}$ within the SSE approach. Consider that we add to the Hamiltonian $\hat{H}$ a linear coupling term $h_B \hat{B}$, i.e., $\hat{H} \to \hat{H} - h_B \hat{B}$, then one obtains for the linear response function

$$\chi_{\hat{A}\hat{B}} = \left. \frac{\partial \langle \hat{A}(h_B) \rangle}{\partial h_B} \right|_{h_B = 0} \quad (38)$$

the Kubo formula

$$\chi_{\hat{A}\hat{B}} = \int_0^\beta d\tau (\langle \hat{A}(\tau) \hat{B}(0) \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle), \quad (39)$$

with $\hat{A}(\tau) = e^{-\tau\hat{H}}\hat{A}e^{\tau\hat{H}}$ the imaginary time-evolved operator. In case $\hat{A}$ and $\hat{B}$ are both diagonal in the local $\hat{S}^z$-basis, this integral can be evaluated using

$$\int_0^\beta d\tau \langle \hat{A}(\tau)\hat{B}(0)\rangle = \left\langle \frac{\beta}{n(n+1)} \left(\sum_{p=0}^{n-1} A(p)\right)\left(\sum_{p=0}^{n-1} B(p)\right) + \frac{\beta}{(n+1)^2}\left(\sum_{p=0}^{n} A(p)B(p)\right)\right\rangle . \tag{40}$$

A special case concerns the uniform magnetic susceptibility, for which $\hat{A} = \hat{B} = \frac{1}{N_s}\sum_{i=1}^{N_s}\hat{S}_i^z$, and thus we obtain the simple expression

$$\chi_u = \frac{\beta}{N_s}\langle \hat{M}_z^2\rangle, \quad \hat{M}_z = \sum_{i=1}^{N_s}\hat{S}_i^z. \tag{41}$$

For further observables, such as the spin stiffness, we refer to Sandvik's review [21]. Let us finally note, that within the world-line approach and the SSE representation it is also possible to evaluate the expectation values of imaginary-time-dependent correlation functions [19, 24]. For an efficient evaluation of the corresponding SSE expressions, based on the explicit mapping between the SSE configuration space and the continuous-time representation, we refer to the review by Assaad and Evertz [25].

3.3 Local updates

In order to sample the SSE configuration space, we need to generate operator-state configurations $(\mathcal{S}_\Lambda, |\sigma\rangle)$ according to the appropriate statistical weights. Updates of the operator sequence can in general not be carried out without affecting also the spin configuration. For example, if one changes the operator at a given propagation level p, say from a diagonal operator $\hat{H}_{1,b}$ on the bond b to the off-diagonal operator $\hat{H}_{2,b}$ on the same bond, then the spin configuration $|\sigma(p)\rangle$ has to change as well. This change in the world-line configuration must be healed at some other propagation level p', by also exchanging diagonal and off-diagonal operators, so that the resulting state $|\sigma(p')\rangle$ is the same as before. Hence, one needs to perform Monte Carlo updates by attempting to change the bond operators at two appropriately picked propagation levels. Indeed, in most cases update attempts with randomly picked positions would not be possible, such that one has to specifically search for operator pairs (or even more operators), which can be updated this way. In the end, such a procedures correspond to the local update schemes that we introduced above for the world-line QMC approach, and they suffer from similar problems.
However, more efficient update schemes have been developed [26, 27]. Such updates in fact consist of two sub-steps: (i) the diagonal update, wherein the expansion order n is modified while keeping fixed the spin configuration $|\sigma\rangle$ of $(\mathcal{S}_\Lambda, |\sigma\rangle)$ as well as the off-diagonal operator content of the operator sequence $\mathcal{S}_\Lambda$, and (ii) a non-diagonal update, the operator-loop update, that updates the operator content as well as the spin configuration simultaneously – similar to, but in a much more efficient way, than in the two-operator case that we discussed above.
We will outline the operator-loop update in the next section, and concentrate here on the diagonal update step, that is an essential local update procedure, which we formulate here for the fixed operator-string length representation [26]. Within this diagonal update, the whole operator sequence is sequentially traversed, and attempts are made to exchange diagonal operators and identity operators, so that in each such step, the expansion order (i.e. the number of non-identity operators) n can change by one to $n \pm 1$. While moving through the operator sequence, the spin configuration $|\sigma\rangle$ is updated whenever an off-diagonal operator is encountered, such that the

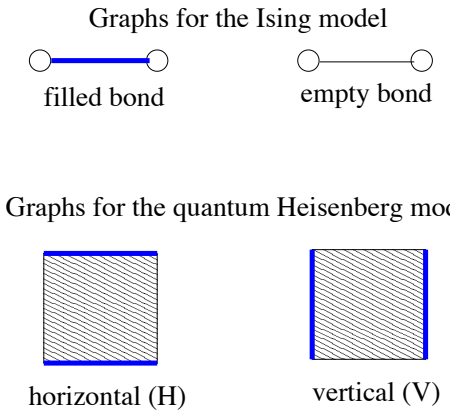


Fig. 5: *Cluster components on a bond vs. graph components on a plaquette (breakups).*

propagated spin configurations $|\sigma(p)\rangle$ are readily available, as they will be required within this update scheme. In more detail, if at propagation level $p = 1, ..., \Lambda$, an identity-operator is present, i.e. $t(p) = 0$, we attempt to replace it by a diagonal bond-operator on a bond b that is randomly chosen among all the N_b lattice bonds. We then use the Metropolis acceptance probability for this move, i.e.

$$P([0,0] \to [1,b]) = \min\left[\frac{\beta N_b}{\Lambda - n}\langle\sigma(p-1)|\hat{H}_{1,b}|\sigma(p-1)\rangle, 1\right], \tag{42}$$

where n is the number of bond-operators in the sequence before the update. For the Heisenberg model considered here, this update can be immediately rejected, if the two spins at the chosen bond at the current propagation level are parallel, while in the other case, the matrix element simply equals $1/2$, cf. Eqs. (28)-(30). On the other hand, if the local operator at the current propagation level is a diagonal bond operator, $t(p) = 1$, an attempt is made to replace it by the identity operator with Metropolis acceptance probability

$$P([1,b] \to [0,0]) = \min\left[\frac{\Lambda - n + 1}{\beta N_b}\langle\sigma(p-1)|\hat{H}_{1,b}|\sigma(p-1)\rangle^{-1}, 1\right]. \tag{43}$$

Once the full operator sequence has been traversed, one should have recovered the initial spin configuration, since $|\sigma(\Lambda)\rangle = |\sigma\rangle$. In case of long-ranged exchange interactions, the above update scheme can be easily generalized to ensure that the various interaction terms are efficiently sampled [28, 29].

4 The loop algorithm

We now present the improved global update schemes that have been developed during the 1990s and early 2000s for the world-line QMC approach. We will first discuss the idea behind the loop algorithm [8], based on the discrete-time formulation. An excellent, general account on the loop algorithm and related methods is provided by Evertz in his review [30].
In order to appreciate the idea behind the loop algorithm, it is useful to first recall the Swendsen-Wang [56] algorithm for the ferromagnetic Ising model. The goal there is to identify physically relevant clusters of parallel spins, and then collectively flip all spins within such a cluster in

order to enhance the update dynamics near criticality. Essentially, one visits all bonds on the lattice and assigns a bond variable $\tau_b \in \{0,1\}$ to each bond, by deciding if a given bond will be filled ($\tau_b = 1$) or kept empty ($\tau_b = 0$) according to certain probabilities, c.f. the upper panel of Fig. 5. Namely, if the two spins connected by the bond are anti-parallel, the bond is kept empty for sure, while for parallel spins along a bond, the bond is filled with a finite probability $p = 1 - e^{-2\beta|J_{ex}|}$, where $J_{ex} < 0$ denotes the strength of the (ferromagnetic) Ising interaction on that bond. Then, one identifies clusters of spins connected by filled bonds, and attempts to flip each cluster individually with probability $1/2$. Within the Wolff-algorithm [57] one instead generates only one of these clusters by starting its construction from a randomly chosen lattice site, and then flips this cluster with probability 1.
For the case of the quantum model in the checkerboard decomposition, we cannot directly apply these schemes of assigning bond variables, since we need to ensure that after an update only valid, unbroken world-line configurations have been generated. Since we are given restrictions for the possible types of shaded plaquettes that can occur in valid configurations, we should thus instead of assigning bond variables assign plaquette variables. Therefore, in the loop algorithm, each plaquette P is assigned a plaquette variable (now called "graph" or "breakup"), with certain probabilities. For the antiferromagnetic Heisenberg model, there are two types of breakups that can be assigned to a plaquette; they are shown in the lower panel of Fig. 5 and we refer to them as the horizontal (H) and vertical (V) breakup. Each shaded plaquette with the state $C|_P$ is assigned the horizontal breakup (H) with probability

$$P(H) = \begin{cases} 0 & , \quad C|_P = \langle\uparrow\uparrow|e^{-\Delta\tau\hat{H}_i}|\uparrow\uparrow\rangle,\ \langle\downarrow\downarrow|e^{-\Delta\tau\hat{H}_i}|\downarrow\downarrow\rangle \\ \tanh(\frac{\Delta\tau J_{ex}}{2}) & , \quad C|_P = \langle\uparrow\downarrow|e^{-\Delta\tau\hat{H}_i}|\uparrow\downarrow\rangle,\ \langle\downarrow\uparrow|e^{-\Delta\tau\hat{H}_i}|\downarrow\uparrow\rangle \\ 1 & , \quad C|_P = \langle\downarrow\uparrow|e^{-\Delta\tau\hat{H}_i}|\uparrow\downarrow\rangle,\ \langle\uparrow\downarrow|e^{-\Delta\tau\hat{H}_i}|\downarrow\uparrow\rangle \end{cases} , \tag{44}$$

otherwise, it is assigned the vertical breakup (V). These graph assignment rules can be obtained upon considering an extended configuration space that combines spin and graph configurations, and which we describe next. Such a general framework to describe cluster algorithms was presented by Kandel and Domany [31], generalizing the Fortuin-Kasteleyn [32] representation of the Ising model. Below we follow the formulation by Kawashima and Gubernatis [33], which makes this relation very transparent. We start from the original representation of the quantum partition function in terms of the spin (world-line) configurations,

$$Z = \sum_C W(C). \tag{45}$$

The phase space is now enlarged, by assigning a set of possible graphs G to the original configurations, such that

$$Z = \sum_C \sum_G W(C,G), \tag{46}$$

where the new weights $W(C,G) \geq 0$ are chosen as to ensure

$$\sum_G W(C,G) = W(C). \tag{47}$$

The algorithm then proceeds as follows: Given a configuration C (which implies $W(C) \neq 0$), we assign first a graph G to the configuration C, chosen with the correct probability,

$$P(G|C) = W(C,G)/W(C). \tag{48}$$

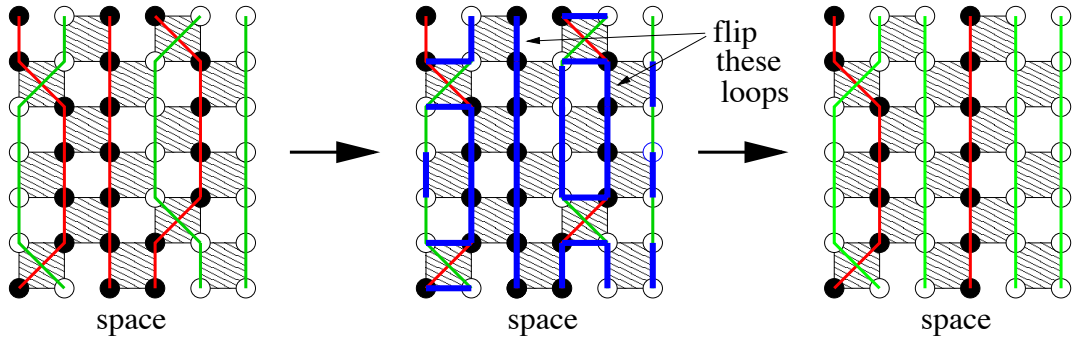


Fig. 6: *An update step in the loop algorithm, where two clusters of spins are flipped.*

Then we choose a new configuration C' with probability $P[(C,G) \to (C',G)]$, keeping the graph G fixed. This completes a configurational update $C \to C'$ and the process repeats by choosing a new graph G', etc. The first step, choosing graphs with probabilities $P(G|C)$, trivially obeys detailed balance. Detailed balance for the second step requires that

$$W(C,G)P[(C,G) \to (C',G)] = W(C',G)P[(C',G) \to (C,G)] \tag{49}$$

One possible solution of this detailed balance condition is provided by the heat-bath algorithm

$$P[(C,G) \to (C',G)] = \frac{W(C,G)}{W(C,G) + W(C',G)}. \tag{50}$$

The whole approach is apparently simplified a lot, if one can find an assignment of graph weights, such that $W(C,G)$ does not depend on the configuration C , whenever it is non-zero in that configuration, i.e., when $W(C,G)$ has the following form:

$$W(C,G) = \Delta(C,G)V(G), \tag{51}$$

where

$$\Delta(C,G) = \begin{cases} 1 & , \quad \text{if } W(C,G) \neq 0 \\ 0 & , \quad \text{otherwise} \end{cases} . \tag{52}$$

In this case, the heat-bath probability (50) simply becomes $P = 1/2$. Furthermore, the weight of the spin configuration $W(C)$ in fact decomposes into a product of plaquette weights, cf. Eq. (17). Further simplifications thus arise, if also the graph weight can be represented as a product of separate weights for each plaquette, i.e., if

$$W(C,G) = \prod_P W_P(C|_P, G|_P), \tag{53}$$

so that the whole graph assignment procedure can be performed locally on the level of the (shaded) plaquettes, i.e. in terms of the plaquette breakups.
While for a more general, anisotropic spin-$1/2$ model additional graph elements are required, the above conditions can be fulfilled for the Heisenberg antiferromagnet by employing only the two breakups (H and V) on each shaded plaquette and assigning the H breakup with the probability given in Eq. (44). An example of such an assignment of the plaquette breakups is shown in the central panel of Fig. 6. After having assigned to each plaquette a breakup, the

graph edges are connected to form clusters of connected graph edges, and hence clusters (loops) of connected local spin states. Now one flips each such cluster independently with probability $P = 1/2$, thereby changing the local spins and thus the world-line configuration. An example of the whole update procedure in shown in Fig. 6. One sees from this example, that after a single loop update, one can realize large-scale changes of the world-line configuration. In particular, as seen in Fig. 6, it is possible to change the total magnetization of the system. The algorithm that we introduced above generates the whole graph (i.e. the complete set of all loops) G and flips each such cluster with probability $P = 1/2$. It is however also possible, to employ a single-loop version of the algorithm, in which, similar to the classical Wolff cluster algorithm, only a single cluster is constructed: One randomly picks a site from the world-line configuration and constructs only the cluster with includes that site. This can be done by determining the breakups (and thus the route of the loop) only on those plaquettes, that are indeed traversed during the loop construction. It is then feasible to employ the Metropolis acceptance for solving the detailed balance condition in Eq. (49), which results in the probability $P = 1$ to flip the just constructed loop. As in the classical case, the single-loop variant usually provides an even smaller dynamical critical scaling than the multi-loop variant.

Furthermore, it is also feasible to realize the loop algorithm directly in the continuous-time version of the world-line QMC formulation, in both multi- and single-loop versions [12]. Since the single-loop version paves a direct conceptual path to the worm, operator- and directed loop algorithms to be introduced below, we focus here on the single-loop variant, even though it is technically a bit more involved than the multi-loop version. The key insight in realizing, that a continuous-time formulation of the loop algorithm is feasible, follows from considering the breakup probability per time, which has a continuous time limit. Indeed, for an imaginary time interval $\tau_1 < \tau_2$, during which the world-line configuration on a given bond does not change, the breakup probability is constant, and the probability density for a horizontal breakup within this imaginary time range becomes

$$\lim_{\Delta\tau\to 0} \frac{P(H)}{\Delta\tau} = \lim_{\Delta\tau\to 0} \frac{\tanh\left(\Delta\tau J_{ex}/2\right)}{\Delta\tau} = \frac{J_{ex}}{2}. \tag{54}$$

We are thus lead to the following procedure: To start the loop construction, a site i is picked randomly, as well as a random imaginary time τ_1 between 0 and β, from which the loop will be constructed, starting (e.g.) initially moving upwards in time. In this forward time direction, one next identifies the time interval $\tau_1 < \tau < \tau_2$, within which the world-line configuration on all sites neighboring site i does not change. For each such neighbor j, draw a random number τ_{ij} from an exponential distribution based on the above probability density for a horizontal breakup, i.e., with $P(\tau_{ij}) \propto \exp(-J_{ex}\tau_{ij}/2)$. Now, let $\bar{j}$ be the neighbor with the smallest value of τ_{ij}, i.e. $\tau_{i\bar{j}} = \min_j(\tau_{ij})$. If $t_d = t_1 + \tau_{i\bar{j}} < t_2$, the loop-end on site i now moves up to the time t_d, and right there jumps to the site $\bar{j}$ (a memorable analogy of this procedure is the radioactive decay with decay constant $J_{ex}/2$ and the neighboring sites corresponding to different decay channels). Since the constructed loops do not self-overlap, one has to exclude in the above procedure all temporal regions of neighboring sites, which have been visited by the current loop already. In case $t_d > t_2$, the loop end stays on site i and moves to time t_2. In case the world-line jumps at time t_2, the loop must jump as well. The whole procedure ensures the constant probability density Eq. (54) for a horizontal breakup. This process is iterated until the loop eventually closes and can be flipped as a whole.

Based on the loop algorithm, large scale simulations of up to a several million quantum spins can be performed even in (quantum) critical regions. After the original introduction of the loop

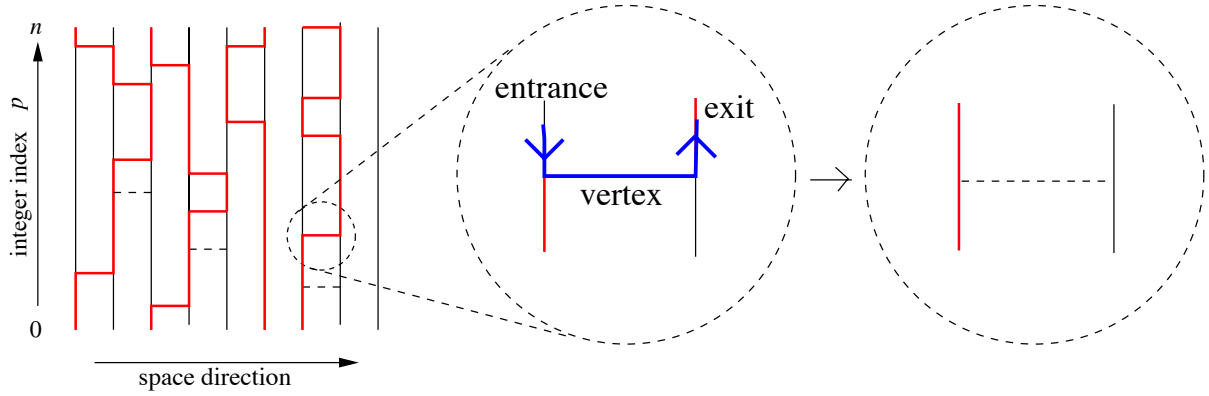


Fig. 7: *A SSE configuration (left panel), the local construction of the operator loop at a 4-leg vertex, corresponding to a bond operator from the SSE operator sequence (central panel), and the resulting final vertex (richt panel). The path taken by the operator loop through the vertex is indicated in the central panel by the arrowed blue line (switch-and-reverse rule).*

algorithm for the spin-1/2 Heisenberg model in a six-vertex model formulation, it has seen various extensions to anisotropic- and higher-spin models, as well as to Hubbard and $t-J$ models. For a detailed derivation of the breakup rules, also beyond the case of the isotropic Heisenberg model considered here, we refer to the review article by Evertz [30]. It is also possible to devise so-called "improved estimators" [34] for certain observables within the loop update, e.g. for correlation functions or magnetic susceptibilities, which provide a significant reduction of statistical errors compared to "naive" estimators for the same quantities. See Evertz's article [30] for details.

Within the SSE formulation, which does not exhibit any systematic (Trotter) error, but still avoids the technical complications of the continuous-time formulation, it is feasible to implement a very simple version of a loop algorithm [26, 35]. This cluster update complements the local diagonal updates, that were introduced in Sec. 3.3. Namely, one considers to construct a discrete-time-like loop, based on a given SSE configuration: Consider a SSE configuration, such as the one shown in Fig. 1(c), redrawn in Fig. 7, as a collection of n vertexes (where n denotes the number of non-identity operators in the operator string), each coming along with 4 "legs". In the central part of Fig. 7, one such a 4-leg-vertex is highlighted. The loop construction starts from a randomly chosen vertex leg in the SSE configuration, called the "entrance leg". Like in the discrete-time version of the loop algorithm, one then defines a local rule of how the loop continues through this vertex and leaves the vertex through an "exit leg". One can interpret such rules as prescribing the scattering of the loop-head off the vertexes. The corresponding probabilities for choosing the exit leg among the four legs of the vertex need to satisfy again a local detailed balance condition. For the antiferromagnetic Heisenberg model it turns out that a very simple, deterministic scattering rule can be devised, namely, that the entrance and exit leg are always located on the same side of the vertex (the switch-and-reverse rule). This is illustrated in the central part of Fig. 7. After the exit leg has been assigned, the loop moves from this exit leg to the leg of another vertex, to which the exit leg is connected within the given SSE configuration, cf. Fig. 7. Now this leg becomes the new entrance leg on the new vertex and the whole process repeats, until the loop closes, i.e., when the last exit leg eventually equals the initial entrance leg on the starting vertex. In the example in Fig. 7, the loop closes when it reaches back to the initial entrance leg after it has visited the vertex atop the vertex from which it

started. All spins along the path of the loops are then flipped. In addition to replacing all visited diagonal operators by off-diagonal operators and vice versa (cf. the right panel of Fig. 7), this operator loop update can also modify the initial spin configuration $|\sigma\rangle$, thus ensuring an ergodic sampling of the SSE configuration space if combined with the diagonal update. Of course, it is also possible in this SSE formulation to construct all the loops for a given SSE configuration and perform a multi-loop update, flipping each such loop with probability $1/2$.

5 Worms, operator loops, and directed loops

The cluster-updates that have been introduced above provide optimal performance for spin inversion symmetric Hamiltonians. However, terms in the Hamiltonian which break this symmetry, such as a magnetic field, are not taken into account during the loop construction. Instead, they would have to be included through the acceptance rate of the loop flips, which can become exponentially small at low temperatures, degrading the algorithm's performance. Hence, it would be highly desirable to generalize the idea of a local construction of a global cluster of spins to cases, where spin-inversion symmetry is not present. The algorithms that we outline next provide such a general sampling scheme.

The worm algorithm [13] formally works in an extended configuration space, where in addition to closed world-line configurations one open world-line fragment (the "worm") is allowed. Formally, this is done by adding a source term to the Hamiltonian,

$$\hat{H}_{\mathrm{worm}} = \hat{H} - \eta \sum_i (\hat{S}_i^+ + \hat{S}_i^-) \; , \tag{55}$$

which allows world lines to be broken with a matrix element proportional to η. The worm algorithm then proceeds as follows: a worm (i.e. a world-line fragment) is created by randomly inserting a pair $(\hat{S}_i^+, \hat{S}_i^-)$ of operators on a world line at nearby times in the world-line configuration. The ends of this worm are then moved randomly in space and time, using local Metropolis or heat bath updates, until the two ends of the worm meet again. Then an update which removes the worm is proposed, and if accepted we are back to a configuration with closed world lines only. This algorithm is straightforward, consisting just of local updates of the worm ends in the extended configuration space but it can perform non-local changes. Furthermore, a worm end can wind around the lattice in the temporal or spatial direction and that way change the magnetization and winding number. While not being as efficient as the loop algorithm in zero magnetic field (the worm movement follows a random walk while the loop algorithm can be interpreted as a self-avoiding random walk), the big advantage of the worm algorithm is that it remains efficient in the presence of a magnetic field. A similar algorithm was actually proposed more than a decade earlier [36]. Instead of a random walk fulfilling detailed balance at every move, the worm head in this earlier algorithm just performed a random walk. The *a posteriori* acceptance rates are then often very small and the algorithm is not efficient, just as the small acceptance rates for loop updates in magnetic fields make the loop algorithm inefficient. This highlights the importance of having the cluster-building rules of a non-local update algorithm closely tied to the physics of the problem.

Algorithms with a similar basic idea are the operator-loop update [26] in the SSE formulation and the directed-loop algorithms [27, 37] which can be formulated in both the SSE and the world-line representation. Like the worm algorithm, these algorithms create two world line discontinuities and move them around by local updates. The operator-loop algorithm for the

SSE representation can be understood as a generalization of the loop algorithm for the SSE configuration, which we described at the end of Sec. 4 [26]. Again, a loop ("operator-loop") is constructed for a given SSE configuration, starting from a random leg of a randomly chosen vertex as the entrance leg. However, for a general Hamiltonian, including e.g. a magnetic field, the scattering rules at the vertices have to be chosen appropriately, and in general turn out to include a stochastic decision, instead of the deterministic switch-and-reverse-rule that holds for the isotropic Heisenberg model. One generic solution of the local detailed balance condition on the scattering rates is provided by a heat-bath choice among the 4 possible exit legs: the probability to chose one out of the four legs is taken to be proportional to the matrix element of the bond operator of the considered vertex in the final resulting spin configuration on this bond [26]. This choice already leads to a rather efficient algorithm. However, it is in general not excluded, that the chosen exit leg is equal to the leg on which the vertex was entered. Such a "bounce" move would result in the operator loop retracing its previous path, and thus in undoing a previously performed change to the spin configuration and the operator content of the operator string. It thus appears desirable to reduce the probability for such bounce moves, or to even eliminate them completely. In fact, for many cases it can be shown that scattering rates can be optimized such as to eliminate completely the bounce move, while still ensuring detailed balance of the operator-loop construction [26, 35, 27, 37, 38]. An example is just the switch-and-reverse-rule for the isotropic Heisenberg antiferromagnet that was presented above. For more general models, a systematic approach to reduce the bounce probability has been formulated [27]. Such algorithms, which direct the loop away from the last changes, are called "directed loop" methods and can be formulated also for the continuous-time world-line approach [27]. Furthermore, it is possible to use linear programming techniques, in order to optimize the scattering rates with respect to the overall bounce probability for general quantum lattice models [38]. An alternative strategy to enhance the update dynamics of the operator loop algorithm is based on the concept of optimal Monte Carlo updates [39, 40]: Starting from the heat-bath solution, one iteratively improves the scattering rules by minimizing its higher eigenvalues in a matrix formulation. Both this "locally optimal solution" and the directed loops are superior to the heat-bath solution and typically perform equally well.

We furthermore want to point out that within the worm and operator-loop approaches, a natural implementation for evaluating transverse spin-spin correlation functions can be realized, based upon on-the-fly measurements of the corresponding matrix elements at the worm-ends, while constructing the worm's path through the world-line configuration [11, 13, 24, 38]. Based on the close connection between the SSE and the continuous-time world-line approach, that we mentioned in Sec. 3.3, it is possible to implement such measurements in a rather efficient way [25]. Finally, we would like to point out that beyond the cluster update methods, which drastically reduce autocorrelation times in QMC simulations of quantum spin systems, it is also possible to adapt various extended ensemble methods such as parallel tempering [41], multicanonical methods [42, 43] and the Wang-Landau technique [44, 45], which are employed in classical Monte Carlo simulations for systems with a rough energy landscape or at first-order phase transitions, to further enhance the performance of QMC simulations in similar situations [46, 47, 48, 23].

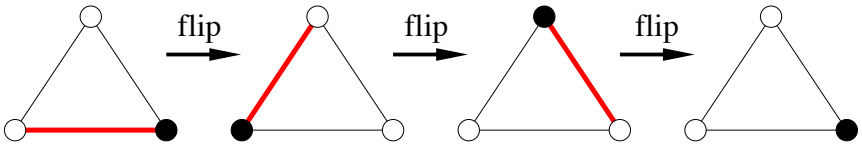


Fig. 8: *Example of an odd number of spin flips on a triangle, returning to the initial state. The bond along which the spins are going to be flipped next is indicated by a red line.*

6 The sign problem

Before closing this short review on world line QMC methods, we finally want to address a severe restriction of this approach: the sign problem. This is indeed its major limitation (note that also other unbiased QMC methods, in particular those for fermionic models, exhibit sign problems in many physically interesting situations) [49]. Currently, it seems rather unlikely that there is a general solution to this problem [50]. Let us explain what the sign problem is and where it comes from. For this purpose, consider again the weight of a plaquette in the Suzuki-Trotter decomposition that corresponds to a local spin exchange:

$$\langle\uparrow\downarrow|e^{-\Delta\tau\hat{H}_i}|\downarrow\uparrow\rangle = -\sinh\left(\frac{\Delta\tau J_{ex}}{2}\right). \tag{56}$$

In the antiferromagnetic case ($J_{ex} > 0$), this matrix element is negative. Then we might actually worry that there will be allowed QMC world line configurations with an overall negative weight, which cannot be assigned as the probability for such a configuration, since probabilities ought to be positive. It is not too hard to convince yourself, that for an open chain or square lattice (in fact any bipartite lattice graph structure) the total number of spin exchange plaquettes in an allowed QMC configuration is even. In that case, the overall weight of the configuration will be positive. Hence, for the Heisenberg antiferromagnet, we are restricted to positive weights only, and can apply the Monte Carlo sampling method of the world-line configurations. However, if the lattice is not bipartite, this property is lost. To see why, consider a single triangle as an example of a non-bipartite lattice. In this case, we can start from one local spin configuration and return to the same configuration, after applying an odd number of (namely 3) spin exchanges, c.f. Fig. 8 for an illustration. Hence, on the triangular lattice, and in fact on any frustrated, non-bipartite lattice, the weight $W(C)$ of a QMC configuration C is not necessarily positive, but can take on negative values as well. How can we then perform a MC sampling of such configurations? Well, at first sight, there seems to be an easy solution: Let us express the partition function of the quantum model in the effective classical representation as follows:

$$Z = \sum_C W(C) = \sum_C |W(C)| \cdot \mathrm{Sgn}(C), \tag{57}$$

where we have written the weight as a product of its absolute value and its sign $\mathrm{Sgn}(C) = W(C)/|W(C)| = \pm 1$. Now, assume, that we would perform a simulation of the antiferromagnet and ignore the sign, thus weighting each configuration according to the absolute values $|W(C)|$ of the weights. In fact, this precisely corresponds to simulating the ferromagnetic Heisenberg model on the same lattice, with the partition function

$$Z_F = \sum_C |W(C)|. \tag{58}$$

Using the generated configurations, we can then obtain the following expectation value of an observable $\hat{A}$ of the antiferromagnetic model,

$$\begin{aligned}\langle\hat{A}\rangle &= \frac{1}{Z}\sum_C W(C)A(C) \\ &= \frac{1}{Z}\frac{Z_F}{Z_F}\sum_C |W(C)|\cdot \mathrm{Sgn}(C)\cdot A(C) \\ &= \frac{Z_F}{Z}\underbrace{\frac{1}{Z_F}\sum_C |W(C)|\cdot \mathrm{Sgn}(C)\cdot A(C)}_{\langle \mathrm{Sgn}\cdot A\rangle_F}, \end{aligned} \tag{59}$$

from (i) the expectation value of $\langle \mathrm{Sgn}\cdot\hat{A}\rangle_F$ measured from the simulation of the ferromagnetic model [i.e., we would measure the expectation value of the product of the sign of a configuration times the value of $\hat{A}$ in that configuration], and (ii) the expectation value of the sign itself, since

$$\frac{Z}{Z_F} = \langle \mathrm{Sgn}\rangle_F, \tag{60}$$

so that

$$\langle\hat{A}\rangle = \frac{\langle \mathrm{Sgn}\cdot\hat{A}\rangle_F}{\langle \mathrm{Sgn}\rangle_F}. \tag{61}$$

Hence, it looks like in order to obtain the expectation value of A for the antiferromagnetic model on a frustrated lattice, we could just simulate the ferromagnetic model [all weights positive], and perform the two measurements of $\langle \mathrm{Sgn}\cdot\hat{A}\rangle_F$ and $\langle \mathrm{Sgn}\rangle_F$. In fact, such an approach works for large temperatures $T \gg J_{ex}$, where the spins do not feel their exchange interactions too much, and behave weakly coupled in both the ferromagnetic and the antiferromagnetic case. But, once we are interested in the low temperature regime, then both models are characterized by very different important states. Can we really expect to learn something about the low temperature behavior of the antiferromagnet from the low temperature behavior of the ferromagnet? No, this should not be reasonable, and indeed, the above approach fails severely at low temperatures. What happens is, that there will be almost the same number of configurations with positive weight as with negative weight, such that the average sign decreases exponentially to zero. Namely, at low temperatures we find

$$\langle \mathrm{Sgn}\rangle_F \sim e^{-\beta N_S \Delta f}, \tag{62}$$

where Δf is the difference in the free energy per site between the antiferromagnetic and the ferromagnetic model (remember, that the free energy $F = -\ln(Z)/\beta$). Hence, in order to calculate the physical observable A from Eq. (61) we divide [in the denominator] by an exponentially small value with a finite statistical error, leading to an exponentially increasing statistical uncertainty for $\langle A\rangle$. Thus, in order to reach low temperatures and large system sizes, the computational time needs to grow exponentially and renders any useful simulation impossible. This behavior is called "the sign problem", and makes QMC simulations of e.g. frustrated quantum magnets practically impossible within the interesting (low-) temperature and large-system regime.

Up to date, no general solution of the QMC sign problem is known, although it can be overcome in certain special cases (cf., e.g., [51, 52, 35]). Moreover, it has been shown that a general

solution to the sign problem essentially constitutes an NP-hard challenge [50]. It is however generally suspected, that no polynomial-time solutions to NP-hard problem exist. Hence, we urge the reader to contact us immediately in case she or he finds a serious path to a generic solution of the QMC sign problem! In the mean time, we hope to have stimulated her or his interest in performing some own QMC simulations for sign-problem free situations. An excellent opportunity to access application-ready codes for the methods that were introduced here is the ALPS (Algorithms and Libraries for Physics Simulations) library [53, 54]. Furthermore, basic SSE simulation codes are also available online [55]. Let the dice roll!

Appendix

A Quantum Heisenberg Model 101

For reference and to fix notation, we define the spin-$1/2$ quantum Heisenberg model on a finite lattice, such as a one-dimensional chain or a two-dimensional square lattice. A spin-$1/2$ quantum spin resides at each lattice site. Each spin $\hat{\mathbf{S}}_i$, located at lattice site $i = 1, \ldots, N_s$ is described by quantum mechanical spin operators

$$\hat{\mathbf{S}}_i = \begin{pmatrix} \hat{S}_i^x \\ \hat{S}_i^y \\ \hat{S}_i^z \end{pmatrix},$$

which fulfill the commutation relation $[\hat{S}_i^\alpha, \hat{S}_j^\beta] = i\hbar\varepsilon_{\alpha\beta\gamma}\hat{S}_i^\gamma\delta_{ij}$, i.e. operators on different sites ($i \neq j$) commute, and locally ($i = j$) the above commutation relations reduce to the well-known algebra of spin operators. In order to set up a basis of the Hilbert space, we start from the eigenvectors of the local $\hat{S}_i^z$ operator at each lattice site i:

$$\begin{aligned} |S_i^z = +1/2\rangle &= |\uparrow\rangle_i = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_i, \\ |S_i^z = -1/2\rangle &= |\downarrow\rangle_i = \begin{pmatrix} 0 \\ 1 \end{pmatrix}_i. \end{aligned}$$

In this local basis, the spin operators are given in matrix form as ($\hbar = 1$),

$$\begin{aligned} \hat{S}_i^z = \frac{1}{2}\sigma_z = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}, \quad \text{i.e.} \quad & \hat{S}_i^z|\uparrow\rangle_i = +\tfrac{1}{2}|\uparrow\rangle_i \\ & \hat{S}_i^z|\downarrow\rangle_i = -\tfrac{1}{2}|\downarrow\rangle_i \\ \hat{S}_i^+ = \hat{S}_i^x + i\hat{S}_i^y = \frac{1}{2}\sigma_x + i\frac{1}{2}\sigma_y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{i.e.} \quad & \hat{S}_i^+|\downarrow\rangle_i = |\uparrow\rangle_i \\ & \hat{S}_i^+|\uparrow\rangle_i = 0 \\ \hat{S}_i^- = \hat{S}_i^x - i\hat{S}_i^y = \frac{1}{2}\sigma_x - i\frac{1}{2}\sigma_y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \text{i.e.} \quad & \hat{S}_i^-|\uparrow\rangle_i = |\downarrow\rangle_i \\ & \hat{S}_i^-|\downarrow\rangle_i = 0 \end{aligned}$$

with the spin raising and lowering operators $\hat{S}_i^+$ and $\hat{S}_i^-$, and the Pauli matrices

$$\begin{aligned} \sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

For a lattice of N_s spins, we have a total of 2^{N_s} basis states of the Hilbert space,

$$\{|S_1^z, ..., S_{N_s}^z\rangle\} = \left\{ \begin{array}{c} |\uparrow,\uparrow,\ldots,\uparrow,\uparrow\rangle \\ |\uparrow,\uparrow,\ldots,\uparrow,\downarrow\rangle \\ |\uparrow,\uparrow,\ldots,\downarrow,\uparrow\rangle \\ \vdots \\ |\uparrow,\downarrow,\ldots,\downarrow,\downarrow\rangle \\ |\downarrow,\downarrow,\ldots,\downarrow,\downarrow\rangle \end{array} \right\}.$$

Each spin operator on site i acts only on the corresponding local spin. Next, we introduce an exchange coupling between nearest neighbor spins, so that the quantum spin system is described by the Hamiltonian

$$\hat{H} = J_{ex} \sum_{\langle i,j \rangle} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j,$$

where $\langle i, j \rangle$ indicates a pair of neighboring lattice sites.
In case $J_{ex} > 0$, the above model is the antiferromagnetic quantum Heisenberg model, which results e.g. within second-order perturbation theory from the half-filled Hubbard model at large local repulsions. The Hamiltonian of the Heisenberg model can also be written using spin raising and lowering operators as

$$\hat{H} = J_{ex} \sum_{\langle i,j \rangle} \hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j = J_{ex} \sum_{\langle i,j \rangle} \left(\frac{1}{2} (\hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+) + \hat{S}_i^z \hat{S}_j^z \right).$$

Consider for example a system of only two spins (a dimer). The Hamiltonian is then given by the matrix

$$\hat{H} = J_{ex} \begin{pmatrix} \frac{1}{4} & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & 0 & \frac{1}{4} \end{pmatrix} \text{ in the basis } \begin{array}{l} |\uparrow\uparrow\rangle \\ |\uparrow\downarrow\rangle \\ |\downarrow\uparrow\rangle \\ |\downarrow\downarrow\rangle \end{array}.$$

References

[1] N. Metropolis, A. R. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, J. of Chem. Phys. **21**, 1087 (1953).

[2] M. Suzuki, Prog. of Theor. Phys. **56**, 1454 (1976).

[3] H.F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959).

[4] M. Suzuki, S. Miyashita and A. Kuroda, Prog. Theor. Phys. **58**, 1377 (1977).

[5] J.E. Hirsch, D.J. Scalapino, R.L. Sugar and R. Blankenbecler, Phys. Rev. Lett. **47**, 1628 (1981).

[6] M.S. Makivić and H.-Q. Ding, Phys. Rev. B **43**, 3562 (1991).

[7] F.F. Assaad and D. Würtz, Phys. Rev. B **44**, 2681 (1991).

[8] H.G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993).

[9] R.M. Fry, Phys. Rev. B **33**, 6271 (1986).

[10] R.M. Fry and R.T. Scalettar, Phys. Rev. B **36**, 3833 (1987).

[11] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Pis'ma v Zh.Eks. Teor. Fiz., **64**, 853 (1996) [English translation: cond-mat/9612091].

[12] B.B. Beard and U.-J. Wiese, Phys. Rev. Lett. **77**, 5130 (1996).

[13] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Sov. Phys. - JETP **87**, 310 (1998).

[14] D.C. Handscomb, Proc. Cambridge Philos. Soc. **58**, 594 (1962).

[15] D.C. Handscomb, Proc. Cambridge Philos. Soc. **60**, 115 (1964).

[16] J.W. Lyklema, Phys. Rev. Lett. **49**, 88 (1982).

[17] D.H. Lee, J.D. Joannopoulos, and J. W. Negele, Phys. Rev. B **30**, 1599 (1984).

[18] E. Manousakis and R. Salvador, Phys. Rev. B **39**, 575 (1989).

[19] A.W. Sandvik and J. Kurkijärvi, Phys. Rev. B **43**, 5950 (1991).

[20] A.W. Sandvik, J. Phys. A **25**, 3667 (1992).

[21] A.W. Sandvik, AIP Conf. Proc. **1297**, 135 (2010).

[22] A. W. Sandvik, R. R. P. Singh, and D. K. Campbell, Phys. Rev. B **56**, 14510 (1997).

[23] S. Wessel, N. Stoop, E. Gull, S. Trebst, and M. Troyer, J. Stat. Mech., P12005 (2007).

[24] A. Dorneich and M. Troyer, Phys. Rev. E **64**, 066701 (2001).

[25] F.F. Assaad and H.G. Evertz, in Springer Lecture Notes in Physics 739, Eds. H. Fehske, R. Schneider and A. Weiße, Springer (2008).

[26] A.W. Sandvik, Phys. Rev. B **59**, R14157 (1999).

[27] O.F. Syljuåsen and A.W. Sandvik, Phys. Rev. E **66**, 046701 (2002).

[28] A.W. Sandvik, Phys. Rev. E **68**, 056701 (2003).

[29] K. Fukui and S. Todo, J. Comput. Phys. **228**, 2629 (2009).

[30] H.G. Evertz, Adv. Phys. **52**, 1 (2003).

[31] D. Kandel and E. Domany, Phys. Rev. B **43**, 8539 (1991).

[32] C. Fortuin and P. Kasteleyn, Physica **57**, 536 (1972).

[33] N. Kawashima and J. Gubernatis, J. Stat. Phys. **80**, 169 (1995).

[34] B. Ammon, H.G. Evertz, N. Kawashima, M. Troyer and B. Frischmuth, Phys. Rev. B **58**, 4304 (1998).

[35] P. Henelius and A.W. Sandvik, Phys. Rev. B **63**, 1102 (2000).

[36] J.J. Cullen and D.P. Landau, Phys. Rev. B **27**, 297 (1983).

[37] O.F. Syljuåsen, Phys. Rev. E **67**, 046701 (2003).

[38] F. Alet, S. Wessel, and M. Troyer Phys. Rev. E **71**, 036706 (2005).

[39] L. Pollet, S.M.A. Rombouts, K. Van Houcke, and K. Heyde, Phys. Rev. E **70**, 056705 (2004).

[40] L. Pollet, K. Van Houcke, and S.M.A. Rombouts, J. Comp. Phys. **225**, 2249 (2007).

[41] E. Marinari and G. Parisi, Europhys. Lett. **19**, 451 (1992).

[42] B.A. Berg and T. Neuhaus, Phys. Lett. B. **267**, 249 (1991).

[43] B.A. Berg and T. Neuhaus, Phys. Rev. Lett. **68**, 9 (1992).

[44] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).

[45] F. Wang and D. P. Landau, Phys. Rev. E **64**, 056101 (2001).

[46] P. Sengupta, A.W. Sandvik, and D.K. Campbell, Phys. Rev. B **65**, 155113 (2002).

[47] M. Troyer, S. Wessel and F. Alet, Phys. Rev. Lett. **90**, 120201 (2003).

[48] M. Troyer, F. Alet, and S. Wessel, Braz. J. of Physics **34**, 377 (2004).

[49] E.Y. Loh Jr., J.E. Gubernatis, R.T. Scalettar, S.R. White, D.J. Scalapino, and R.L. Sugar, Phys. Rev. B **41**, 9301 (1990).

[50] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).

[51] T. Nakamura, Phys. Rev. B **57**, R3197 (1998).

[52] S. Chandrasekharan and U.-J. Wiese, Phys. Rev. Lett. **83**, 3116 (1999).

[53] A. F. Albuquerque, F. Alet, P. Corboz, P. Dayal, A. Feiguin, S. Fuchs, L. Gamper, E. Gull, S. Gürtler, A. Honecker, R. Igarashi, M. Körner, A. Kozhevnikov, A. Läuchli, S. R. Manmana, M. Matsumoto, I. P. McCulloch, F. Michel, R. M. Noack, G. Pawłowski, L. Pollet, T. Pruschke, U. Schollwöck, S. Todo, S. Trebst, M. Troyer, P. Werner, and S. Wessel, J. Magn. Mater. **310**, 1187 (2007).

[54] Codes available online at http://alps.comp-phys.org.

[55] Codes available online at http://physics.bu.edu/∼sandvik.

[56] R.H. Swendsen and J.S. Wang, Phys. Rev. Lett **58**, 86 (1987).

[57] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).

B 5 Dynamics of open quantum systems [1]

H. Schoeller
Institut für Theorie der Statistischen Physik
RWTH Aachen

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

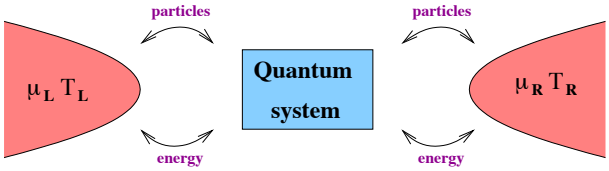


Fig. 1: *A small quantum system coupled to several infinitely large reservoirs via energy and/or particle exchange. The reservoirs are characterized by temperatures T_α and chemical potentials μ_α.*

1 Introduction

An open quantum system consists of a local quantummechanical system of fixed size coupled to infinitely large reservoirs in statistical equilibrium via a well-defined interaction, see Fig. 1 for a sketch of the system. The analysis of the time evolution of the reduced density matrix of the local system is of fundamental importance for nonequilibrium statistical mechanics. Of particular interest is the interplay of quantum coherence in the local quantum system and dissipation generated by the reservoirs, which is the reason why this field is called *dissipative quantum mechancis* . Many topics are here of current interest: (1) The development of a microscopic theory for irreversible time evolution of the local density matrix; (2) The characterization of the reduced dynamics, in particular by generic features independent of the microscopic details of high-energy processes; (3) The calculation of typical relaxation and decoherence rates; (4) The analysis of quantum fluctuations induced by the system-reservoir interaction beyond perturbation theory; (5) The analysis of the influence of strong correlations in the local quantum system, induced by charging energies; (6) The analysis of the influence of inhomogeneous boundary conditions, induced by different temperatures and/or chemical potentials of several reservoirs, leading to energy, particle, and spin currents; (7) The analysis of deviations of the stationary local density matrix from a grandcanonical one, induced by quantum fluctuations from the system-reservoir coupling or by the presence of several reservoirs; (8) The development of nonequilibrium renormalization group (RG) methods capable of resumming logarithmic divergencies occuring in higher-order perturbation theory in the system-reservoir coupling, either at high energies w.r.t. the band width of the reservoirs or at low energies w.r.t. the inverse time in the long-time limit; (9) The study of non-Markovian dynamics leading to additional terms in the time evolution with unexpected oscillation frequencies and decay rates together with non-exponential time evolution; (10) The crossover between coherent and incoherent dynamics induced by the sign and size of the system-reservoir coupling and other tunable parameters. It is the purpose of this tutorial introduction to present a microscopic theory for the time evolution of open quantum systems, to discuss some of the above aspects from a generic point of view, and to characterize explicitly the reduced dynamics for elementary 2-level quantum systems coupled via energy, particle or spin exchange to external reservoirs.

Although the field of dissipative quantum mechanics has a long history, the field has regained an enormous interest in the last decades due to its relevance in quantum transport phenomena in nanoelectronic systems and quantum information processing, and due to the controlled realization of low-dimensional quantum systems in cold atom gases. To describe the time evolution of the reduced density matrix of the local quantum system microscopically, one starts from the von

Neumann equation for the total density matrix $\rho_{\rm tot}(t)$ of the full system (we set $\hbar = e = k = 1$)

$$i\,\dot{\rho}_{\rm tot}(t) \;=\; [H_{\rm tot}(t), \rho_{\rm tot}(t)] \;=\; L_{\rm tot}(t)\,\rho_{\rm tot}(t) \quad , \tag{1}$$

where $L_{\rm tot}(t)$ is the so-called Liouville operator, a superoperator which acts on an arbitrary operator via $L_{\rm tot}(t)A = [H_{\rm tot}(t), A]$. The central idea is always to integrate out the reservoir degrees of freedom and to set up a formally exact kinetic equation for the local density matrix $\rho(t) = {\rm Tr}_{\rm res}\rho_{\rm tot}(t)$, defined by the trace ${\rm Tr}_{\rm res}$ over the reservoir degrees of freedom of the total density matrix. This kinetic equation has the form

$$i\,\dot{\rho}(t) \;=\; \int_{t_0}^{t} dt'\, L(t,t')\,\rho(t') \quad , \tag{2}$$

where t_0 is the initial time and $L(t,t')$ is an effective Liouville operator acting only on operators of the local quantum system. This superoperator contains all the information of the reservoir degrees of freedom and the system-reservoir interaction. For a time-translational invariant Hamiltonian, $L(t,t') = L(t-t')$ depends only on the relative time difference. The effective Liouvillian $L(t,t')$ is only defined for $t > t'$, i.e. it acts as a response function relating the density matrix at time t' to the one at the later time t. This acounts for memory effects and leads to non-Markovian dynamics. The only assumption needed to derive the kinetic equation (2) is the factorization of the total density matrix at the initial time t_0 in an arbitrary local part $\rho(t_0)$ and an equilibrium part for the reservoirs

$$\rho_{\rm tot}(t_0) \;=\; \rho(t_0)\,\rho_{\rm res}^{\rm eq} \quad , \quad \rho_{\rm res}^{\rm eq} \;=\; \prod_{\alpha} \rho_{\alpha}^{\rm eq} \quad , \quad \rho_{\alpha}^{\rm eq} \;=\; \frac{1}{Z_\alpha}\, e^{-(H_\alpha - \mu_\alpha N_\alpha)/T_\alpha} \quad , \tag{3}$$

where T_α, μ_α, H_α, N_α, and Z_α are the temperature, the chemical potential, the Hamiltonian, the particle number, and the partition function of reservoir α, respectively. However, by changing the Hamiltonian at a certain quench time $t_q > t_0$ abruptly, other initial conditions can be realized where system and reservoirs are correlated.

Various techniques have been developed to calculate the effective Liouvillian $L(t,t')$. The traditional ones are projection operator techniques [1] and functional integrals [2]. Recently, a quantum field theoretical approach has been developed, which allows for a systematic classification of all processes in all orders of perturbation theory in the system-reservoir coupling [3]. With this method, it is possible to go beyond bare perturbation theory which is necessary at low temperatures due to various logarithmic divergencies at high and low energies. The method is capable of identifying these logarithmic divergencies very effectively and an RG method in nonequilibrium has been set up to resum them. This allows a systematic weak-coupling expansion in the renormalized coupling constants to be formulated with which the time evolution on all time scales even when the reservoirs have different chemical potentials or temperatures can be discussed. This technique has been applied successfully to the Kondo model [4, 5, 6], the interacting resonant level model (IRLM) [7, 8, 9], and the ohmic spin boson model [10]. In particular, it has turned out that the RG formulation is most effective for the calculation of the time evolution if the Fourier variable E conjugate to the time t is used as flow parameter, i.e. as the paramater w.r.t. which derivatives of the various quantities of interest are taken to obtain differential equations (the so-called RG equations). This technique is called the E-RTRG method [6, 10]. The models treated so far fall into the special class where the density of states

in the reservoirs and the frequency dependence of the system-reservoir couplings is weak and varies on the scale of the high-energy cutoff D. Physically, the high-energy cutoff can either be the band width of the reservoirs or it is some internal high-energy scale of the local quantum system, like e.g. charging energies, arising when effective models are used by integrating out high-energy processes (e.g. quantum dots in the Coulomb blockade regime where charge degrees of freedom can be eliminated, see the lecture B3 by T. Costi). For such models it is often possible to find universal physics where the special form of the high-energy cutoff function is not important and influences only the value of certain low-energy scales (e.g. the Kondo temperature for the Kondo model). In such a case the high-energy cutoff D does no longer occur explicitly. Furthermore, for a wide class of time-translational invariant models it turns out that the effective Liouvillian has the form

$$L(E) \;=\; L_\Delta(E) \;+\; E\,L'(E) \quad , \tag{4}$$

where $L(E) = \int_0^\infty dt e^{iEt} L(t)$ is the Fourier-transform of the response function $L(t-t') = L(t,t')\theta(t-t')$. In this decomposition $L_\Delta(E)$ and $L'(E)$ are slowly varying logarithmic functions, where $L_\Delta(E)$ is proportional to some energy scale Δ of the model which can be anything except for the Fourier variable E. This form will be shown by the RG analysis in Section 5 for the concrete models under consideration but it remains an interesting question for the future how generic this form is. A large part of this tutorial deals with the technical details of calculating the appearing functions L_Δ and L'. Before we do that, we will first investigate the physical consequences for the time evolution in Section 3. We will see that when $L(E)$ has the form (4) the time evolution can generically be decomposed as

$$\rho(t) \;=\; \sum_n F_n(t)\, e^{-iz_n t}\, \rho_{t=0} \quad , \tag{5}$$

where $z_n = \pm\Omega_n - i\Gamma_n$, with $\Omega_n, \Gamma_n \geq 0$, determine the oscillation frequencies and decay rates of exponential decay, and $F_n(t)$ are pre-exponential functions, which typically consist of power-laws $\sim 1/t^k$ ($k = 1, 2, \dots$) and logarithmic corrections in the long-time limit $t \gg 1/|z_n|$. At least one of the exponential scales is zero $z_{\rm st} = 0$, which determines the stationary state.

It is the purpose of the present article to first discuss the generic physics of the time evolution on the basis of the form (4) of the effective Liouvillian, and with this motivation discuss the E-RTRG method for the calculation of $L(E)$ and its decomposition into (4). Then we will summarize the results for the time evolution of the Kondo model, the ohmic spin boson model, and the IRLM. We note that other RG methods have been developed recently to discuss the time evolution of open quantum systems. The most important ones are the flow-equation method [11] and the functional RG [12]. The latter will be introduced in the lecture B7 by V. Meden and is a method where one expands systematically in the short-ranged renormalized interaction parameter present in the local system but not in the system-reservoir coupling, i.e. it is complementary to the RTRG technique where arbitrary local interactions can be treated but an expansion in the renormalized system-reservoir coupling is needed. Besides the analytical RG methods, there is also an extensive research going on to develop numerical methods to describe the time evolution, like e.g. time-dependent numerical renormalization group [13], time-dependent density matrix renormalization group [14], iterative stochastic path integrals [15], and quantum Monte Carlo [16]. Furthermore, for special models, field-theoretical methods have been used to find exact results [17].

2 Basic models

We start with the description of the basic models under consideration, where the quantum system consists of 2 states coupled via spin (Kondo model), charge and potential (IRLM), or energy (spin boson) fluctuations to a noninteracting environment. The total Hamiltonian is assumed to be time-translational invariant and consists of three parts

$$H_{\rm tot} \;=\; H \;+\; H_{\rm res} \;+\; V \quad , \quad H_{\rm res} \;=\; \sum_{\alpha} H_{\alpha} \quad , \quad H_{\alpha} \;=\; \sum_{k\sigma} \epsilon_{\alpha\sigma k}\, a^{\dagger}_{\alpha\sigma k} a_{\alpha\sigma k} \tag{6}$$

where H is the Hamiltonian of the local quantum system, V is the system-reservoir interaction, and $H_{\rm res}$ describes the noninteracting (fermionic or bosonic) reservoirs with field operators $a_{\alpha\sigma k}$. α is the reservoir index, σ the channel index (e.g. spin), and the quantum number k characterizes the energy. For convenience, for given α and σ, we will denote by $\omega = \epsilon_{\alpha\sigma k} - \mu_{\alpha}$ the energy of the reservoir states measured relative to the chemical potential, and we assume that the relation between ω and k is unique. As a consequence, the field operators of the reservoirs can be characterized by the multi-index $1 \equiv \eta\alpha\sigma\omega$, where $\eta = \pm$ distinguishes between creation ($\eta = +$) and annihilation operators ($\eta = -$). Depending on the model under consideration, we will define below convenient forms of the field operators $a_1 \equiv a_{\eta\alpha\sigma}(\omega)$ in continuum notation, such that the commutation relations read (the upper/lower case refers always to bosons/fermions)

$$[a_{\alpha\sigma}(\omega), a^{\dagger}_{\alpha'\sigma'}(\omega')]_{\mp} \;=\; \delta_{\alpha\alpha'}\, \delta_{\sigma\sigma'}\, \delta(\omega - \omega')\, \rho_{\alpha\sigma}(\omega) \quad , \tag{7}$$

where $[\cdot,\cdot]_{\mp}$ denotes the commutator/anticommutator for bosons/fermions. As defined below the spectral function $\rho_{\alpha\sigma}(\omega)$ contains the d.o.s. of the reservoirs and possibly frequency-dependencies of the system-reservoir couplings. Together with the commutation relations

$$[a_1, H_{\alpha}] \;=\; -\eta\,(\omega \,+\, \mu_{\alpha})\, a_1 \quad , \quad [a_1, N_{\alpha}] \;=\; -\eta\, a_1 \quad , \tag{8}$$

it follows that the contraction of two reservoir field operators w.r.t. the equilibrium distribution is given by

$$\overbrace{a_1\, a_{1'}} \;=\; {\rm Tr}_{\rm res}\, a_1\, a_{1'}\, \rho^{\rm eq}_{\rm res} \;=\; \delta_{1\bar{1}'}\, \rho_{\alpha\sigma}(\omega)\, f^{\eta}_{\alpha}(\omega) \;=\; \delta_{1\bar{1}'} \left\{ \begin{array}{c} \eta \\ 1 \end{array} \right\} \rho_{\alpha\sigma}(\omega)\, f_{\alpha}(\eta\omega) \quad , \tag{9}$$

where $\bar{1} \equiv -\eta\alpha\sigma\omega$ is defined by reversing the sign of η, $\delta_{12} = \delta_{\eta_1\eta_2}\delta_{\alpha_1\alpha_2}\delta_{\sigma_1\sigma_2}\delta(\omega_1 - \omega_2)$, $f^{+}_{\alpha}(\omega) = f_{\alpha}(\omega)$, $f^{-}_{\alpha}(\omega) = 1 \pm f_{\alpha}(\omega)$, and $f_{\alpha}(\omega) = (e^{\omega/T_{\alpha}} \mp 1)^{-1}$ is the Bose/Fermi distribution.

In terms of the continuum field operators, the system-reservoir interaction V is generically written as a sum of terms of the form

$$V \;=\; \frac{1}{n!} \left\{ \begin{array}{c} 1 \\ \eta_1\eta_2 \ldots \eta_n \end{array} \right\} :\, a_n a_{n-1} \ldots a_1 \,:\, g_{12\ldots n} \;\rightarrow\; \frac{1}{n!}\, g_{12\ldots n} \;:\, a_1 a_2 \ldots a_n \,: \quad , \tag{10}$$

where $n = 1, 2, \ldots$ is any integer, imlicit summation/integration is assumed over the multi-indices $i \equiv \eta_i\alpha_i\sigma_i\omega_i$, the operator $g_{12\ldots n}$ acts only on the local system, and $: \cdots :$ denotes normal-ordering w.r.t. to the equilibrium distribution (3) of the reservoirs (i.e. in any Wick-decomposition contractions are not allowed within the normal-ordered expression). We call the operators $g_{1\ldots n}$ n-point vertex operators since, together with the corresponding superoperators (55), they will appear in the diagrammatic technique as vertices with n reservoir lines, see

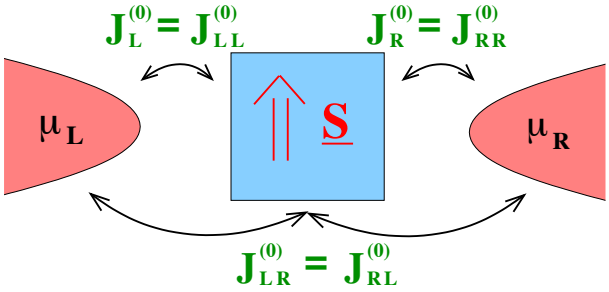


Fig. 2: *A sketch of the nonequilibrium Kondo model. A local spin is coupled via isotropic exchange couplings $J^{(0)}_{\alpha\alpha'}$ to the reservoir spins. The two reservoirs are characterized by the same temperature T but the chemical potentials μ_L and μ_R can be different defining the voltage $V = \mu_L - \mu_R$ across the system. The nondiagonal exchange couplings $J^{(0)}_{LR} = J^{(0)}_{RL}$ describe spin exchange processes where a particle is transferred between the reservoirs, giving rise to a current.*

Section 4. For bosons the two forms for V shown in (10) are the same. For fermions, the first form is needed for the definition of the vertex operators and, for n odd, $g_{1\ldots n}$ is of fermionic nature and anticommutes with the reservoir field operators. However, it can be shown [3] that, for the calculation of any local observables, the second form for V can be used and local and reservoir operators can be taken as if they commute. The vertex operators have the properties

$$g_{1\ldots i\ldots j\ldots n} \;=\; \pm g_{1\ldots j\ldots i\ldots n} \quad , \quad g^\dagger_{1\ldots n} \;=\; g_{\bar n\ldots\bar 1} \quad . \tag{11}$$

The first relation can always be achieved by (anti-)symmetrization of the reservoir field operators within the normal-ordering in (10), whereas the second one is necessary for the property $V = V^\dagger$. In the following we will specify the definition of the continuum reservoir field operators a_1, the spectral density $\rho_{\alpha\sigma}(\omega)$ and the vertex operators $g_{1\ldots n}$ for the various models.

The Kondo model. In its most basic form the Kondo model describes a local spin-$\frac{1}{2}$ system coupled via short-ranged and isotropic exchange couplings to fermionic reservoir spins, see Fig. 2 for a sketch of the system. It is a model system to describe local spin fluctuations. For the case of a single channel the Hamiltonian reads

$$H \;=\; h^{(0)}\, S_z \quad , \quad V \;=\; \sum_{\alpha\alpha'} \frac{J^{(0)}_{\alpha\alpha'}}{\sqrt{\rho^{(0)}_\alpha \rho^{(0)}_{\alpha'}}} \sum_{\sigma\sigma'} \frac{1}{2}\underline{\sigma}_{\sigma\sigma'} \,:\, \sum_k a^\dagger_{\alpha\sigma k} \sum_{k'} a_{\alpha'\sigma' k'} \,:\, \underline{S} \quad , \tag{12}$$

where the isotropic exchange couplings $J^{(0)}_{\alpha\alpha'} = J^{(0)}_{\alpha'\alpha}$ are real and dimensionless, $\underline{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices, $\underline{S}$ is the local spin, and $h^{(0)}$ is the local bare magnetic field. $\rho^{(0)}_\alpha$ denotes the d.o.s. of the reservoirs at the Fermi level. The Kondo model can be derived via a Schrieffer-Wolff transformation from the single-impurity Anderson model (see the lecture B3 by T. Costi), in which case the exchange couplings fulfil the relation

$$J^{(0)}_{\alpha\alpha'} \;=\; 2\sqrt{x_\alpha x_{\alpha'}}\, J^{(0)} \quad , \quad \sum_\alpha x_\alpha \;=\; 1 \quad , \quad 0 < x_\alpha < 1 \quad , \tag{13}$$

where x_α are asymmetry factors weighting the energy broadening of the local level from reservoir α. Defining the continuum field operator by $a_1 = \frac{1}{\sqrt{\rho^{(0)}_\alpha}} \sum_k \delta(\omega - \epsilon_{\alpha\sigma k} + \mu_\alpha) a_{\eta\alpha\sigma k}$, with

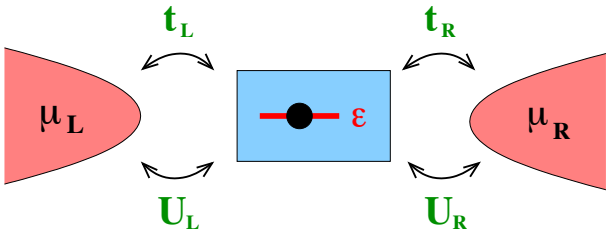


Fig. 3: *A sketch of the IRLM. A local level without spin is coupled via tunneling and Coulomb interaction to several reservoirs.*

$1 \equiv \eta\alpha\sigma\omega$, and assuming a flat d.o.s. in the reservoirs, we obtain for the spectral function and the vertex operator

$$\rho_{\alpha\sigma}(\omega) \;=\; \rho(\omega) \;=\; \frac{1}{\rho_\alpha^{(0)}} \sum_k \delta(\omega - \epsilon_{\alpha\sigma k} + \mu_\alpha) \;=\; \frac{D^2}{D^2+\omega^2} \quad , \tag{14}$$

$$g_{11'} \;=\; \frac{1}{2}\, J^{(0)}_{\alpha\alpha'}\, \underline{\sigma}_{\sigma\sigma'} \cdot \underline{S} \quad , \quad \text{for} \quad \eta = -\eta' = + \quad , \tag{15}$$

where D is the band width of the reservoirs and, for convenience, we have chosen a Lorentzian for the high-energy cutoff function. The case $\eta = -\eta' = -$ is obtained from $g_{11'} = -g_{1'1}$.

The IRLM. The IRLM is a basic model to describe charge and potential fluctuations. It consists of a single fermionic level, which is coupled to fermionic reservoirs via tunneling and a local Coulomb interaction, see Fig. 3 for a sketch of the system. Disregarding the spin, the Hamiltonian is defined by

$$H \;=\; \epsilon\, c^\dagger c \quad , \quad V \;=\; \frac{t_\alpha}{\sqrt{\rho_\alpha^{(0)}}} \sum_k \left(a^\dagger_{\alpha k} c \;+\; \text{h.c.} \right) \;+\; \frac{U_\alpha}{\rho_\alpha^{(0)}} \left(c^\dagger c - \frac{1}{2} \right) \sum_{kk'} :\, a^\dagger_{\alpha k} a_{\alpha k'} \,: \quad , \tag{16}$$

where c is the field operator annihilating a particle on the local system, t_α are the tunneling matrix elements (in units of $\frac{1}{\sqrt{\rho_\alpha^{(0)}}}$), U_α denote the dimensionless Coulomb couplings, and ϵ is the bare energy of the local level. At $\epsilon = 0$ the model fulfils particle-hole symmetry. Defining the continuum field operators as for the Kondo model, with $1 \equiv \eta\alpha\omega$ (i.e. omitting the spin index), we find the same result (14) for the spectral function, and the vertex operators are given by

$$g_1 \;=\; t_\alpha \begin{cases} c & \text{for } \eta = + \\ c^\dagger & \text{for } \eta = - \end{cases} \quad , \quad g_{11'} \;=\; \delta_{\eta,-\eta'}\, \delta_{\alpha\alpha'}\, \eta\, U_\alpha \left(c^\dagger c - \frac{1}{2} \right) \quad . \tag{17}$$

The spin boson model. The spin boson model describes energy fluctuations, where a 2-level system is coupled linearly to a phonon bath, see Fig. 4 for a sketch of the system. The Hamiltonian is given by

$$H_{\text{res}} \;=\; \sum_k \omega_k\, a^\dagger_k a_k \quad , \quad H \;=\; \frac{1}{2}\, \epsilon\, \sigma_z \;-\; \frac{1}{2}\, \Delta\, \sigma_x \quad , \quad V \;=\; \frac{1}{2}\, \sigma_z \sum_k \alpha_k \left(a_k \;+\; a^\dagger_k \right) \quad , \tag{18}$$

where ϵ and Δ denote the bias and the tunneling of the local 2-level system, respectively. The phonon frequencies $\omega_k > 0$ are positive, and the equilibrium phonon distribution is character-

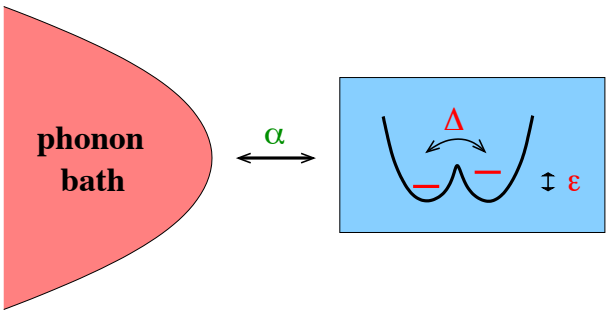


Fig. 4: *A sketch of the spin boson model. A 2-level system, characterized by tunneling* Δ *and bias* ϵ *is coupled linearly via the dimensionless coupling constant* α *to a phonon bath of harmonic oscillators.*

ized by temperature T. The k-dependence of the real coupling constants α_k and the phonon frequencies ω_k is considered by defining the continuum field operators by $a_1 = \sum_k \alpha_k \delta(\omega - \omega_k) a_{\eta k}$ with $1 \equiv \eta\omega$. This leads to the following spectral function and vertex operator

$$\rho(\omega) \;=\; \sum_k \alpha_k^2\, \delta(\omega - \omega_k) \;=\; 2\,\alpha\,\omega\, \left(\frac{\omega}{D}\right)^{s-1} \theta(\omega)\, \frac{D^2}{D^2+\omega^2} \quad , \quad g_1 \;=\; \frac{1}{2}\,\sigma_z \quad , \tag{19}$$

where α is a dimensionless coupling constant, and we have again chosen a Lorentzian high-energy cutoff function with band width D. The special form chosen for $\rho(\omega)$ describes the ohmic case for $s = 1$ considered in this article, whereas $s < 1$ ($s > 1$) define the sub-ohmic (super-ohmic) cases. For the special case $\Delta = 0$ the spin boson model can be solved exactly [2] with the result

$$\langle \sigma_{x,y} \rangle(t) \;=\; e^{-h(t)}\, \langle \sigma_{x,y} \rangle_{t=0} \quad , \quad \langle \sigma_z \rangle(t) \;=\; \langle \sigma_z \rangle_{t=0} \quad , \tag{20}$$

with $h(t) = -\int d\omega (\rho(\omega)/\omega)(1-\cos(\omega t))(1+2f(\omega))$, where $f(\omega) = (exp(\omega/T)-1)^{-1}$ is the Bose function.

For the special case of α close to $\frac{1}{2}$, the ohmic spin boson model can be mapped on the IRLM with a single reservoir (with $\mu = 0$) [2]. The parameters U and t of the IRLM are related to α and Δ of the ohmic spin boson model in the following way

$$U \;=\; 1 \,-\, \sqrt{2\alpha} \quad , \quad \Gamma^{(0)} \;=\; 2\,\pi\, t^2 \;=\; \frac{\Delta^2}{D} \quad . \tag{21}$$

The local occupation $\langle n \rangle(t) = \langle c^\dagger c \rangle(t)$ of the IRLM is related to the expectation value $\langle \sigma_z \rangle(t)$ of the ohmic spin boson model via

$$2\,\langle n \rangle(t) \,-\, 1 \;=\; \langle \sigma_z \rangle(t) \quad , \tag{22}$$

whereas the expectation value $\langle \sigma_{x,y} \rangle(t)$ of the spin boson model is related to expectation values of highly nonlinear operators involving reservoir degrees of freedom in the IRLM. The value $\alpha = \frac{1}{2}$ is of special importance since, at this point, the time evolution of $\langle \sigma_z \rangle(t)$ changes from an oscillating one (for $\alpha < \frac{1}{2}$) to a purely decaying one (for $\alpha > \frac{1}{2}$) [2, 8, 9]. Correspondingly, for the IRLM, this crossover occurs when the sign of the Coulomb interaction U is changed.

3 Kinetic equation and time evolution

In this section we aim at discussing the time evolution from a generic point of view based on the general form (2) of the kinetic equation and the form (4) of the effective Liouvillian for the case of a time-translational invariant Hamiltonian. Using $L(t,t') = L(t-t')$ the kinetic equation reads

$$i\dot{\rho}(t) \;=\; \int_0^t dt'\, L(t-t')\,\rho(t') \quad , \tag{23}$$

where, for convenience, we have set the initial time $t_0 = 0$. The reduced density matrix $\rho(t)$ acts only in local space, i.e. has matrix elements $\rho(t)_{ss'} = \langle s|\rho(t)|s'\rangle$, where s and s' are states of the local quantum system. In contrast, the superoperator $L(t)$ acts on local operators A, i.e. the matrix elements can be written as $L_{s_1s_2,s_1's_2'} = \langle s_1 s_2|L(t)|s_1' s_2'\rangle$, where $|ss'\rangle = |s\rangle\langle s'|$ are the basis elements (= operators) in Liouville space and $\langle ss'|A = \langle s|A|s'\rangle$ are the corresponding dual vectors. The density matrix fulfils the property of conservation of probability $\mathrm{Tr}\rho(t) = 1$ and is self-adjoint $\rho(t) = \rho(t)^\dagger$. It is straightforward to show that the kinetic equation respects these properties if and only if the effective Liouville operator fulfils the properties

$$\mathrm{Tr}\, L(t) \;=\; \sum_s L(t)_{ss,\cdot\cdot} \;=\; 0 \quad , \quad L(t)^c \;=\; -\,L(t) \quad , \tag{24}$$

where the c-transform is defined by $L(t)^c_{s_1s_2,s_1's_2'} = L(t)^*_{s_2s_1,s_2's_1'}$ and fulfils the useful property $(L(t)A)^\dagger = L(t)^c A^\dagger$. In Fourier space $L(E) = \int_0^\infty dt e^{iEt} L(t)$ this means

$$\mathrm{Tr}\, L(E) \;=\; 0 \quad , \quad L(E)^c \;=\; -\,L(-E^*) \quad , \tag{25}$$

or for the quantities $L_\Delta(E)$ and $L'(E)$ appearing in the decomposition (4)

$$\mathrm{Tr}\, L_\Delta(E) \;=\; \mathrm{Tr}\, L'(E) \;=\; 0 \quad , \quad L_\Delta(E)^c \;=\; -L_\Delta(-E^*) \quad , \quad L'(E)^c \;=\; L'(-E^*) \quad . \tag{26}$$

With $\rho(E) = \int_0^\infty dt e^{iEt}\rho(t)$, the kinetic equation reads in Fourier space $E\rho(E) - i\rho_{t=0} = L(E)\rho(E)$ leading to the formal solution

$$\rho(E) \;=\; i\,R(E)\,\rho_{t=0} \quad , \quad R(E) \;=\; \frac{1}{E \,-\, L(E)} \quad . \tag{27}$$

We now investigate the consequences of the generic form (4) of the effective Liouvillian $L(E)$. Using inverse Fourier transform, the time evolution can be calculated for $t > 0$ from

$$\rho(t) \;=\; \frac{i}{2\pi} \int_{-\infty+i0^+}^{\infty+i0^+} dE\, e^{-iEt}\, R(E)\,\rho_{t=0} \;=\; \frac{i}{2\pi} \int_{-\infty+i0^+}^{\infty+i0^+} dE\, e^{-iEt}\, \tilde{R}(E)\, Z'(E)\,\rho_{t=0} \quad , \tag{28}$$

where we have defined

$$\tilde{R}(E) \;=\; \frac{1}{E \,-\, \tilde{L}_\Delta(E)} \quad , \quad \tilde{L}_\Delta(E) \;=\; Z'(E)\,L_\Delta(E) \quad , \quad Z'(E) \;=\; \frac{1}{1 \,-\, L'(E)} \quad . \tag{29}$$

By convention, $Z'(E)$ is called the Z'-factor operator. The last form of (28) is very helpful for the evaluation of the energy integral because it explicitly exhibits the slowly varying logarithmic

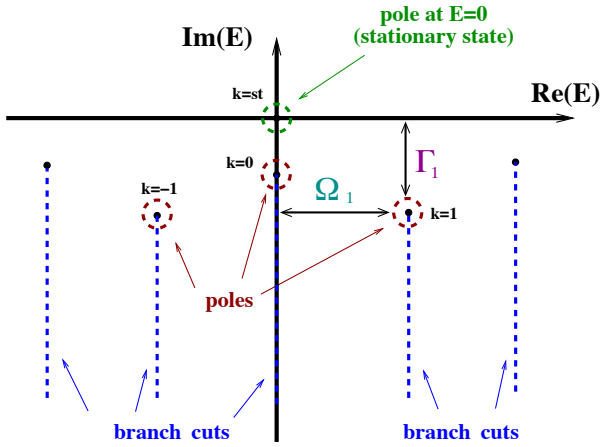


Fig. 5: *The analytic structure of the resolvent $R(E) = 1/(E - L(E))$. Generically, there is a pole at $E = z^p_{st} = 0$ corresponding to the stationary state. Branch cuts can occur in the lower half starting either at a branching point or at a pole. This will be demonstrated in Section 4, see Eq.* (88). *The analytic structure is symmetric w.r.t. the imaginary axis. The poles are denoted by $z^p_k = \pm\Omega_k - i\Gamma_k$. A pole lying on the imaginary axis gets the index $k = 0$, the others are labelled by $\pm k$ with $k = 1, 2, \ldots$.*

functions $\tilde{L}_\Delta(E)$ and $Z'(E)$. The energy integral $\int dE$ is calculated by closing the integration contour in the lower half of the complex plane and deforming the contour such that the poles and branch cuts of the integrand are enclosed, see Fig. 5. To identify the singularities of the integrand we use the spectral decomposition of the Liouvillian $\tilde{L}_\Delta(E)$ in terms of its eigenvalues $\lambda_k(E)$ and corresponding projectors $P_k(E)$

$$\tilde{L}_\Delta(E) \;=\; \sum_k \lambda_k(E)\, P_k(E) \quad . \tag{30}$$

Since we deal with a non-hermitian superoperator, we have to distinguish the left and right eigenvectors, which we denote in Dirac notation by $|x_k(E)\rangle$ and $\langle\bar{x}_k(E)|$

$$\tilde{L}_\Delta(E)\,|x_k(E)\rangle \;=\; \lambda_k(E)\,|x_k(E)\rangle \quad , \quad \langle\bar{x}_k(E)|\,\tilde{L}_\Delta(E) \;=\; \langle\bar{x}_k(E)|\,\lambda_k(E) \quad . \tag{31}$$

The eigenvectors fulfil the orthonormalization condition $\langle\bar{x}_k(E)|x_{k'}(E)\rangle = \delta_{kk'}$ and the projectors are given by $P_k(E) = |x_k(E)\rangle\langle\bar{x}_k(E)|$ with $\sum_k P_k(E) = 1$.

Due to the condition $\mathrm{Tr}\tilde{L}_\Delta(E) = 0$, we obtain either $\lambda_k(E) = 0$ or $\mathrm{Tr}\,|x_k(E)\rangle = 0$. Therefore, the Liouvillian has always an eigenvalue zero, which we characterize by the index $k = \mathrm{st}$ since it corresponds to the stationary state (see below). The other eigenvalues are numerated by $k = 0, \pm1, \pm2, \ldots$. Normalizing the eigenvector with $k = \mathrm{st}$ according to $\mathrm{Tr}|x_{\mathrm{st}}(E)\rangle = 1$ and using $\langle\bar{x}_{\mathrm{st}}(E)| = \mathrm{Tr}$, we get

$$\mathrm{Tr}\,|x_{\mathrm{st}}(E)\rangle \;=\; \sum_s \langle ss|x_{\mathrm{st}}(E)\rangle \;=\; 1 \quad , \quad \langle\bar{x}_{\mathrm{st}}(E)|ss\rangle \;=\; 1 \tag{32}$$

$$\mathrm{Tr}\,|x_k(E)\rangle \;=\; \sum_s \langle ss|x_k(E)\rangle \;=\; 0 \quad , \quad \text{for} \quad k = 0, \pm1, \pm2, \ldots \quad . \tag{33}$$

As a consequence we get

$$P_{\mathrm{st}}(E) \;=\; |x_{\mathrm{st}}(E)\rangle\,\mathrm{Tr} \quad , \tag{34}$$

and the property $\mathrm{Tr} L_\Delta(E) = \mathrm{Tr} L'(E) = 0$ can also be written as

$$P_{\text{st}}(E)\, Z'(E) \;=\; P_{\text{st}}(E) \quad , \quad P_{\text{st}}(E)\, L_\Delta(E) \;=\; 0 \quad . \tag{35}$$

Due to the condition $\tilde{L}_\Delta(E)^c = -\tilde{L}_\Delta(-E^*)$, the eigenvalues and projectors occur always in pairs (except for $k = 0$, st where we define $k \equiv -k$) with

$$\lambda_{-k}(E) \;=\; -\lambda_k(-E^*) \quad , \quad P_{-k}(E) \;=\; P_k(-E^*)^c \quad . \tag{36}$$

Using the spectral representation, the time evolution can be written as

$$\rho(t) \;=\; \frac{i}{2\pi} \sum_k \int_\gamma dE\, e^{-iEt}\, \frac{1}{E - \lambda_k(E)}\, P_k(E)\, Z'(E)\, \rho_{t=0} \quad , \tag{37}$$

where γ is an integration contour which encloses the lower half of the complex plane including the real axis. Poles are located at $E = z_k^p = \lambda_k(z_k^p) = \pm\Omega_k - i\Gamma_k$, with $\Omega_k, \Gamma_k \geq 0$, where $z_{\text{st}}^p = 0$ is a pole at the origin, see Fig. 5. At **zero temperature**, which we consider from now on, additional nonanalytic features occur from branch cuts since $\lambda_k(E)$, $P_k(E)$ and $Z'(E)$ depend logarithmically via terms $\sim \ln(\frac{D}{E - z_n})$ generated by the ultraviolet divergencies from the high-energy cutoff D (at finite temperature the branch cuts turn into an infinite number of discrete poles separated by $2\pi T$). From the structure of the perturbation theory (see below) it can be seen that the singularities z_n are associated with poles of the resolvents $\tilde{R}(E_{1\ldots n})$, where

$$E_{1\ldots n} \;=\; E + \bar{\mu}_{1\ldots n} \quad , \quad \bar{\mu}_{1\ldots n} \;=\; \bar{\mu}_1 + \cdots + \bar{\mu}_n \quad , \quad \bar{\mu}_1 \;=\; \eta_1\, \mu_{\alpha_1} \quad , \tag{38}$$

i.e. are located at $z_n = E$ with $E_{1\ldots n} = E + \bar{\mu}_{1\ldots n} = z_k^p$. Therefore, the singularities $z_n = z_k^p - \bar{\mu}_{1\ldots n}$ are generically given by the poles shifted by some linear combination of the chemical potentials of the reservoirs.

In Section 5 we will see how $\tilde{L}_\Delta(E)$ and $Z'(E)$ can be determined from differential equations, see Eq. (111), where we differentiate w.r.t the Fourier variable E. These differential equations are defined in the whole complex plane and will be the RG equations of the E-RTRG method. E is called the flow paramater and a solution of the RG equations along a certain path is called the RG flow. The particular advantage is that these RG equations can be solved along the paths $E = z_n + i\Lambda \pm O^+$, with Λ real, starting at some high value $\Lambda \sim D$ down to $\Lambda = -\infty$. Since no singularities are present on these paths, it can even be numerically enforced that the branch cuts start at z_n and point into the direction of the negative imaginary axis. Furthermore, the jump of the Liouvillian at the branch cuts can be determined from the difference of the two solutions and the integrals around the branch cuts can be calculated. The choice that the branch cuts point into the direction of the negative imaginary axis is very convenient since $e^{-iEt} = e^{-iz_n t} e^{-xt}$ is exponentially decaying in xt, which allows an analytical discussion of the long-time limit (see below). Using $E = z_n - ix \pm O^+$, the integration around a particular branch cut (including the case when the branching point is a pole) gives the contribution $\rho_n(t) = F_n(t) e^{-iz_n t} \rho_{t=0}$ to the time evolution with

$$F_n(t) \;=\; \frac{1}{2\pi} \int_0^\infty dx\, e^{-xt} \left\{ R(z_n - ix + 0^+) - R(z_n - ix - 0^+) \right\} \tag{39}$$

such that the total time evolution can be written in the form (5)

$$\rho(t) \;=\; \sum_n \rho_n(t) \;=\; \sum_n F_n(t)\, e^{-iz_n t}\, \rho_{t=0} \quad . \tag{40}$$

For the further evaluation of $F_n(t)$ it is important to distinguish between the cases when the branching point is a pole or not. We label the contributions from branching poles z_k^p by $F_k^p(t)$ and $\rho_k^p(t)$ and the others by $F_n^b(t)$ and $\rho_n^b(t)$, such that (40) reads

$$\begin{aligned}\rho(t) \;&=\; \sum_k \rho_k^p(t) \;+\; \sum_n \rho_n^b(t) \\ &=\; \sum_k F_k^p(t)\, e^{-iz_k^p t}\, \rho_{t=0} \;+\; \sum_n F_n^b(t)\, e^{-iz_n^b t}\, \rho_{t=0} \quad .\end{aligned} \tag{41}$$

Thereby we note that the same singularity $z_k^p = z_n^b$ can appear as a branching pole and as a branching point, since a certain term involving $\lambda_k(E)$ in (37) can have a branch cut at $z_n^b = z_{k'}^p$ with $k' \neq k$. Generically, for weakly coupled system-reservoir systems, the contributions $\rho_n^b(t)$ are smaller since they are proportional to the system-reservoir coupling (see below Eq. (88)). However, if the decay rates occuring in z_n^b are smaller than those ones of z_k^p, the relative order of the various terms can change as function of time, as discussed e.g. in detail in Refs. [8, 9] for the IRLM with positive Coulomb interaction or the ohmic spin boson model for α close but slightly below the value $\alpha = \frac{1}{2}$. In the Markovian approximation, only the contributions $\rho_k^p(t)$ remain and the pre-exponential functions are approximated by constants of $O(1)$.

Time-evolution regimes. Using the general expressions (28) and (37), one can discuss the qualitative form of the time evolution in different time regimes. For **short times** $t \ll 1/|z_n|$, only high frequencies $E \sim 1/t \gg |z_n|$ matter in Eq. (28), i.e. the cutoff scales z_n in the logarithmic terms are unimportant and can be neglected. Furthermore, to leading order, we can replace $E \to 1/t$ in the logarithmic parts, and we obtain from (28)

$$\rho(t) \;=\; \frac{i}{2\pi} \int_\gamma dE\, e^{-iEt}\, \frac{1}{E - \tilde{L}_\Delta(1/t)}\, Z'(1/t)\, \rho_{t=0} \;=\; e^{-i\tilde{L}_\Delta(1/t)t}\, Z'(1/t)\, \rho_{t=0} \quad . \tag{42}$$

Expanding the exponential one finds in leading order that the logarithmic dependence of $Z'(1/t)$ and $\tilde{L}_\Delta(1/t)$ at high energies determine the short time behavior. This means that the RG equations are cut off at the large energy scale $E = 1/t$, which is the poor man scaling regime, where all the cutoff scales z_n are unimportant. In this regime the time evolution is determined by the scaling of $Z'(1/t)$ and $\tilde{L}_\Delta(1/t)$. If, in addition, $t \gg 1/D$, where D is the high-energy cutoff, one obtains universal time evolution in the short-time regime. It means that all leading logarithmic divergencies $\sim (\alpha \ln(Dt))^n$ have been resummed in the functions $Z'(1/t)$ and $\tilde{L}_\Delta(1/t)$, where $\alpha \ll 1$ is some small dimensionless coupling parameter. Based on this unified picture the universal short-time behaviour has been derived in Refs. [5, 9, 10] for the Kondo model, the IRLM, and the ohmic spin boson model, in accordance with similiar results of previous literature.

For **intermediate and long times** $t \gtrsim 1/|z_n|$, we have to study the contributions from the poles and branch cuts in detail, based on the decomposition (41). We start with the contributions from the **branch cuts starting at a pole** z_k^p, which we evaluate by using the form (37). For the branch cut integral we set $E = z_k^p - ix \pm 0^+$ and replace in leading order $\lambda_k(E) \to z_k^p$ and the

logarithmic function $P_k(E)Z'(E)$ by its average $\bar{P}_k(z_k^p - ix)\bar{Z}'(z_k^p - ix)$ over the branch cut, where $\bar{A}(E) = \frac{1}{2}(A(E+0^+)+A(E-0^+))$. Furthermore, in leading order, we can use $x \to 1/t$ in the logarithmic functions. This gives the result

$$F_k^p(t) \;\approx\; \frac{1}{2\pi}\int_{0^-}^{\infty} dx\, e^{-xt}\left(\frac{1}{-ix+0^+} - \frac{1}{-ix-0^+}\right)\bar{P}_k(z_k^p - i/t)\,\bar{Z}'(z_k^p - i/t) \quad . \tag{43}$$

Using $\frac{1}{-ix+0^+} - \frac{1}{-ix-0^+} = 2\pi\delta(x)$, we obtain the following contribution to the total time evolution (40)

$$\rho_k^p(t) \;\approx\; \bar{P}_k(z_k^p - i/t)\,\bar{Z}'(z_k^p - i/t)\,e^{-iz_k^p t}\,\rho_{t=0} \quad , \tag{44}$$

i.e., for $z_k^p = \pm\Omega_k - i\Gamma_k$, an exponential one with oscillation Ω_k and decay rate Γ_k, modulated by a logarithmic function. For the special term $k = \mathrm{st}$, where $z_{\mathrm{st}}^p = 0$, $P_{\mathrm{st}}(E) = |x_{\mathrm{st}}(E)\rangle\mathrm{Tr}$ and $P_{\mathrm{st}}(E)Z'(E) = P_{\mathrm{st}}(E)$, we get the following contribution to the time evolution

$$\rho_{\mathrm{st}}^p(t) \;\approx\; |\overline{x_{\mathrm{st}}}(-i/t)\rangle \;\xrightarrow{t\to\infty}\; \rho_{\mathrm{st}} \;=\; |\overline{x_{\mathrm{st}}}(0)\rangle \quad , \tag{45}$$

i.e. we see that for $t \to \infty$ one always gets the stationary distribution ρ_{st} but, if z_{st}^p is a branching pole, logarithmic corrections can occur for the time evolution which do not decay exponentially. We note that for the models discussed here, there is no logarithmic term in the diagrammatic series involving the pole z_{st}^p. In addition, there is no accidental pole $z_{k\neq\mathrm{st}}^p = 0$, and therefore the pole at $E = 0$ is isolated and has no attached branch cuts.

The evaluation of a **branch cut starting at a branching point** z_n^b which is not a pole is more subtle since both $\lambda_k(E)$ and $P_k(E)Z'(E)$ can be discontinuous and cancellations can occur between the two contributions. Therefore, it is more convenient to start from the first expression of (28) involving the resolvent $R(E)$. Denoting by $\delta A = A_+ - A_-$ the jump across the branch and by $\bar{A} = \frac{1}{2}(A_+ + A_-)$ the average value, with $A_\pm = A(E \pm 0^+) = \bar{A} \pm \frac{1}{2}\delta A$, one finds for the jump of the resolvent expanding in the small quantity δL (leading to higher orders in the renormalized coupling constants)

$$\delta R(E) \;=\; R_+\,\delta L\,R_- \;=\; \frac{1}{E-\bar{L}}\,\delta L\,\frac{1}{E-\bar{L}} \;+\; O(\delta L^3) \quad . \tag{46}$$

Using $\overline{AB} - \bar{A}\bar{B} = \frac{1}{4}\delta A\delta B$, we get

$$\frac{1}{E-\bar{L}} \;=\; \overline{\frac{1}{E-L}} + O(\delta L^2) \;=\; \sum_k \overline{\frac{1}{E-\lambda_k}P_kZ'} + O(\delta L^2) \;=\; \sum_k \frac{1}{E-\bar{\lambda}_k}\bar{P}_k\bar{Z}' + O(\delta L^2) \tag{47}$$

Inserting this in (46), neglecting $O(\delta L^3)$, and approximating $E = z_n^b - ix \to z_n^b - i/t$ in the logarithmic functions $\bar{\lambda}_k$, $\bar{P}_k$ and $\bar{Z}'$, we get the following result for the branch cut integral

$$F_n^b(t) \;\approx\; \frac{1}{2\pi}\sum_{\substack{kk' \\ z_k^p, z_{k'}^p \neq z_n^b}} \int_0^\infty dx\, e^{-xt}\,\frac{1}{z_n^b - ix - \bar{\lambda}_k^n}\,\bar{P}_k^n\,\bar{Z}'^n\,\delta L(z_n^b - ix)\,\frac{1}{z_n^b - ix - \bar{\lambda}_{k'}^n}\,\bar{P}_{k'}^n\,\bar{Z}'^n\,, \tag{48}$$

where $\bar{\lambda}_k^n = \bar{\lambda}_k(z_n^b - i/t)$, $\bar{P}_k^n = \bar{P}_k(z_n^b - i/t)$ and $\bar{Z}'^n = \bar{Z}'(z_n^b - i/t)$. We have omitted the cases $z_k^p = z_n^b$ or $z_{k'}^p = z_n^b$ since we consider a branching point and not a branching pole. Since $\bar{\lambda}_k^n \approx z_k^p$, we can neglect x in the denominators of the resolvents for times $t \sim 1/x \gg$

$1/|z_n^b - z_{k,k'}^p|$. In this case, the long-time scaling is determined by the scaling of $\delta L(z_n^b - ix)$ for small x. Besides additional logarithmic corrections (which again can be treated by replacing $x \to 1/t$), we will show in Section 5 that

$$\delta L(z_n^b - ix) \sim \theta(x) \tag{49}$$

for models with charge fluctuations (like the IRLM) and

$$\delta L(z_n^b - ix) \sim x\,\theta(x) \tag{50}$$

for models with spin/orbital or energy fluctuations (like the Kondo and the ohmic spin boson model), see Eq. (120). Therefore, if x can be neglected in the resolvents of the integrand of (48), we obtain (up to logarithmic corrections) $\rho_n^b(t) \sim 1/t$ for charge fluctuations and $\rho_n^b(t) \sim 1/t^2$ for spin/orbital and energy fluctuations. For special resonant cases, where z_n^b comes close to z_k^p or $z_{k'}^p$, one can also define time regimes $1/|z_n^b| \lesssim t \ll 1/|z_n^b - z_{k,k'}^p|$, where x dominates in the denominators for certain values of k or k', leading to different scaling. If x is not neglected in (88), the integral can also be calculated exactly, leading typically to exponential integrals from which the whole crossover behaviour from intermediate $t \sim 1/|z_n|$ to long times $t \gg 1/|z_n|$ can be calculated.

In the regime of intermediate to long times the cutoff scales z_n are very important. Each term of the series (41) has to be treated separately, leading to different scaling of the individual terms (in contrast to the short-time regime, where all exponentials can be approximated by one and only the sum of all pre-exponential functions matters). As we have seen above, various functions $K(z_n - i/t)$ with logarithmic scaling occur in the projectors, the Z-factors, and the jump of the Liouvillian. In bare perturbation theory, the logarithmic functions $K(E)$ will contain powers of terms $\sim \alpha \ln \frac{D}{E - z_m}$. To get rid of the high-energy cutoff D, a standard technique is to resum first all leading logarithmic divergencies $\sim (\alpha \ln \frac{D}{\Lambda_c})^n$, where $\Lambda_c \gtrsim |z_m|$ is some maximal physical low energy scale. Technically, this can be achieved by cutting off the RG flow at Λ_c, defining renormalized coupling constants α_c at this point, and expanding the full solution for $|E| \lesssim |z_m|$ in α_c. This is possible if α_c is small, i.e. if Λ_c is much larger than the strong coupling scale Λ^*, where the coupling constants become of $O(1)$. As a result, $K(E)$ will contain powers of logarithmic terms $\sim \alpha_c \ln \frac{\Lambda_c}{E - z_m}$. For $E = z_n - i/t$ the most dangerous case is $n = m$, leading to powers in the time-dependent parameter $\alpha_t \sim \alpha_c \ln \Lambda_c t$. Since $\alpha_c \ll 1$, this parameter is small $\alpha_t \ll 1$, unless time is exponentially large. Therefore, it can be treated perturbatively, leading to logarithmic corrections $\sim \alpha_t$ in the pre-exponential functions. This strategy has been used in Refs. [5, 7, 8, 9, 10] to determine the time evolution at intermediate and long times (but not exponenitally large times) for the Kondo model, the IRLM, and the spin boson model.

Finally, the most complicated time regime is the one at **exponentially large times**, where $\alpha_t \sim O(1)$. In this regime, a perturbative treatment is no longer possible and all powers of α_t are important. These logarithmic divergencies at *low energies* are independent of those at *large energies* and can even arise if there is no logarithmic divergence at high energies. Their occurence is related to the fact that, concerning the time evolution, the final cutoff scale at low energies is set by inverse time $1/t$ and not by decay rates. The latter holds only for the calculation of stationary properties, see Refs. [3, 4, 18]. The E-RTRG method is unique in the sense that it is also capable of resumming the logarithmic divergencies at low energies, provided the renormalized coupling constants remain small when E approaches one of the singularities

z_n. Recently, this has been achieved in a controlled way for the ohmic spin boson model [10], where deviations from previously predicted scaling behaviour have been found. Results for the Kondo model and the IRLM are still under investigation in this regime. In particular for the Kondo model, the problem is that the renormalized coupling constants become of $O(1)$ when approaching one of the singularities although they might be small for the calculation of stationary quantities at $E = 0$. Thus, weak-coupling problems for stationary quantities can turn into strong-coupling ones for the calculation of the long-time behaviour at exponentially large times.

4 Diagrammatic expansion

Effective Liouvillian. In this section we will derive a quantum field theoretical diagrammatic representation of the effective Liouvillian by expanding in the system-reservoir interaction V, following Refs. [3, 19]. Although this can be done for the general case of an explicitly time-dependent Hamiltonian [9], here we will restrict ourselves to the more simpler case of a time-translational invariant Hamiltonian. To find a diagrammatic expansion of the effective Liouvillian $L(E)$ in Fourier space, we try to bring the local density matrix $\rho(E)$ into the form (27). We start from the formal solution of the von Neumann equation (1) for the total density matrix, use the initial condition (3), and obtain by expanding in the system-reservoir interaction

$$\begin{aligned}
\rho(E) \;&=\; \int_0^\infty dt\, e^{iEt}\, \mathrm{Tr}_{\mathrm{res}}\, \rho_{\mathrm{tot}}(t) \;=\; \int_0^\infty dt\, e^{iEt}\, \mathrm{Tr}_{\mathrm{res}}\, e^{-iL_{\mathrm{tot}}t}\, \rho_{\mathrm{tot}}(t=0) \\
&=\; \mathrm{Tr}_{\mathrm{res}}\, \frac{i}{E \,-\, L_{\mathrm{tot}}}\, \rho_{t=0}\, \rho_{\mathrm{res}}^{(\mathrm{eq})} \;=\; \mathrm{Tr}_{\mathrm{res}}\, \frac{i}{E \,-\, L^{(0)} \,-\, L_{\mathrm{res}} \,-\, L_V}\, \rho_{t=0}\, \rho_{\mathrm{res}}^{(\mathrm{eq})} \\
&=\; i\, \mathrm{Tr}_{\mathrm{res}}\, R^{(0)}(E - L_{\mathrm{res}}) \sum_{k=0}^{\infty} (L_V\, R^{(0)}(E - L_{\mathrm{res}}))^k\, \rho_{t=0}\, \rho_{\mathrm{res}}^{(\mathrm{eq})} \quad , \qquad (51)
\end{aligned}$$

where we have defined

$$R^{(0)}(E) \;=\; \frac{1}{E \,-\, L^{(0)}} \quad , \quad L^{(0)} \;=\; [H, \cdot] \quad , \quad L_{\mathrm{res}} \;=\; [H_{\mathrm{res}}, \cdot] \quad , \quad L_V \;=\; [V, \cdot] \quad . \qquad (52)$$

Using the form (10) of the system-reservoir interaction, a similiar form can be derived for the Liouville superoperator L_V

$$L_V \;=\; \frac{1}{n!} \sum_{p=\pm} G_{1\ldots n}^{(0)p\ldots p} \;:\, A_1^p \ldots A_n^p \,: \quad . \qquad (53)$$

Here, $p = \pm$ is the so-called Keldysh index, which indicates whether the interaction V arises from the first or the second part of the commutator $L_V b = Vb - bV$ (b is an arbitrary operator). A_1^p are reservoir field superoperators in Liouville space defined by

$$A_1^p\, b \;=\; \sigma_{\mathrm{res}}^p \left\{ \begin{array}{ll} a_1\, b & \text{for } p = + \\ b\, a_1 & \text{for } p = - \end{array} \right. \quad , \qquad (54)$$

and $G_{1\ldots n}^{(0)p\ldots p}$ is a superoperator acting in Liouville space of the local quantum system defined by

$$G_{1\ldots n}^{(0)p\ldots p}\, b \;=\; \left\{ \begin{array}{ll} 1 & \text{for } n \text{ even} \\ \sigma^p & \text{for } n \text{ odd} \end{array} \right\} \left\{ \begin{array}{ll} g_{1\ldots n}\, b & \text{for } p = + \\ -b\, g_{1\ldots n} & \text{for } p = - \end{array} \right. \quad . \qquad (55)$$

σ^p and $\sigma^p_{\rm res}$ are convenient sign superoperators which account for fermionic signs and measure the parity of the fermionic particle number difference $N_s - N_{s'}$ of intermediate states $|ss'\rangle$ in Liouville states via the definition (N_s denotes the particle number of state s and $\pm$ refers to bosons/fermions)

$$\sigma^+ \,=\, 1 \quad , \quad \sigma^-_{s_1 s_2, s_1' s_2'} \,=\, \delta_{s_1 s_1'}\, \delta_{s_2 s_2'}\, (\pm)^{N_s - N_{s'}} \quad , \tag{56}$$

and a corresponding definition for $\sigma^p_{\rm res}$ by replacing local states s by reservoir states. Since the total parity (local system plus reservoirs) of all intermediate states must be even in Liouville space for fermions (note that it is impossible to prepare a nondiagonal matrix element of the total density matrix where the total fermionic particle number difference is odd, see Refs. [19, 20] for a detailed discussion and the consequences of this point), we obtain the important property

$$\sigma^p\, \sigma^p_{\rm res} \,=\, 1 \quad . \tag{57}$$

From the definition of the reservoir field superoperators one can straightforwardly derive how the product $: A^p_1 \ldots A^p_n :$ occuring in Eq. (53) acts in Liouville space

$$: A^p_1 \ldots A^p_n :\, b \,=\, \left\{ \begin{array}{ll} 1 & \text{for } n \text{ even} \\ \sigma^p_{\rm res} & \text{for } n \text{ odd} \end{array} \right\} \left\{ \begin{array}{ll} : a_1 \ldots a_n :\, b & \text{for } p = + \\ b \,: a_1 \ldots a_n : & \text{for } p = - \end{array} \right. \quad , \tag{58}$$

i.e. similiar to $G^{(0)p\ldots p}_{1\ldots n}$ but a minus sign is missing for $p = -$. Taking this equation together with (55) and using the property (57), one can easily prove the representation (53) for L_V.

Most importantly, the reservoir field superoperators are defined such that the usual Wick theorem can be applied (see Ref. [19] for an elegant proof), i.e. the average $\mathrm{Tr}_{\rm res} A^{p_1}_1 \ldots A^{p_n}_n \rho_{\rm res}$ decomposes into a product of pair contractions and the sum has to be taken over all combinations, with the usual definition of a fermionic sign to disentangle the various contractions. Using (9) a single contraction is given by the expression

$$\begin{aligned} \gamma^{pp'}_{11'} \,&=\, \overbrace{A^p_1\, A^{p'}_{1'}} \,=\, \mathrm{Tr}_{\rm res}\, A^p_1\, A^{p'}_{1'}\, \rho^{\rm eq}_{\rm res} \\ &=\, \delta_{1\bar{1}'} \left\{ \begin{array}{c} 1 \\ p' \end{array} \right\} \rho_{\alpha\sigma}(\omega)\, f^{p'\eta}_\alpha(\omega) \,=\, \delta_{1\bar{1}'}\, p' \left\{ \begin{array}{c} \eta \\ 1 \end{array} \right\} \rho_{\alpha\sigma}(\omega)\, f_\alpha(p'\eta\omega) \quad . \end{aligned} \tag{59}$$

Using the form (53) in (51) one can shift all reservoir field superoperators $A^{p_i}_i$ to the right by using the analog of the commutation relation (7) in Liouville space

$$A^p_1\, L_{\rm res} \,=\, (L_{\rm res} \,-\, \eta(\omega + \mu_\alpha))\, A^p_1 \quad . \tag{60}$$

This means that by shifting a certain field superoperator A^p_1 through all resolvents to the right, we shift all reservoir Liouville operators $L_{\rm res}$ standing right to A^p_1 by $-\eta(\omega + \mu_\alpha)$, where $1 \equiv \eta\alpha\sigma\omega$. We note that, with the second form (10) of the interaction, there is no fermionic sign when commuting local and reservoir operators. Shifting all reservoir field superoperators to the right and using the notation

$$\begin{aligned} X_{1\ldots n} \,=\, \bar{\omega}_{1\ldots n} \,+\, \bar{\mu}_{1\ldots n} \quad , \quad \bar{\omega}_{1\ldots n} \,&=\, \bar{\omega}_1 \,+\, \ldots\, \bar{\omega}_n \quad , \quad \bar{\omega}_1 \,=\, \eta_1\, \omega_1 \quad , \\ \bar{\mu}_{1\ldots n} \,&=\, \bar{\mu}_1 \,+\, \ldots\, \bar{\mu}_n \quad , \quad \bar{\mu}_1 \,=\, \eta_1\, \mu_1 \quad , \end{aligned} \tag{61}$$

we obtain for (51) the form

$$\begin{aligned}\rho(E) \;=\; i\sum_{k=0}^{\infty} \mathrm{Tr}_{\mathrm{res}}\, R^{(0)}(E-L_{\mathrm{res}})\,(\frac{1}{n!}G^{(0)})\,R^{(0)}(E+X_{M_1}-L_{\mathrm{res}})\cdot \\ \cdot\,(\frac{1}{n!}G^{(0)})\,R^{(0)}(E+X_{M_2}-L_{\mathrm{res}})\,\ldots\,(\frac{1}{n!}G^{(0)})\,R^{(0)}(E+X_{M_k}-L_{\mathrm{res}})\,\rho_{t=0}\cdot \\ \cdot\,(:A\ldots A:)\,(:A\ldots A:)\,\ldots\,(:A\ldots A:)\,\rho_{\mathrm{res}}^{(\mathrm{eq})} \\ =\; i\sum_{k=0}^{\infty} R^{(0)}(E)\,(\frac{1}{n!}G^{(0)})\,R^{(0)}(E+X_{M_1})\,\ldots\,(\frac{1}{n!}G^{(0)})\,R^{(0)}(E+X_{M_k})\,\rho_{t=0}\cdot \\ \cdot\,\mathrm{Tr}_{\mathrm{res}}\,\{(:A\ldots A:)\,\ldots\,(:A\ldots A:)\}\,\rho_{\mathrm{res}}^{(\mathrm{eq})} \quad , \end{aligned} \tag{62}$$

where we have used $\mathrm{Tr}_{\mathrm{res}} L_{\mathrm{res}} = 0$ in the last step. Thereby, the set M_i includes those indices of reservoir field superoperators which were standing left to the corresponding resolvent in the original expression. As a result the local and reservoir degrees of freedom have been decoupled and the trace over the reservoir degrees of freedom can be performed by the application of Wick's theorem in Liouville space. Since all diagrams give the same contribution when the indices of a particular vertex $G^{(0)p\ldots p}_{1\ldots n}$ are permuted, the factor $\frac{1}{n!}$ is cancelled, except for the case when two vertices are connected by m contractions, leaving a symmetry factor $\frac{1}{m!}$. This leads to a sum of diagrams which symbolically are translated by the rule

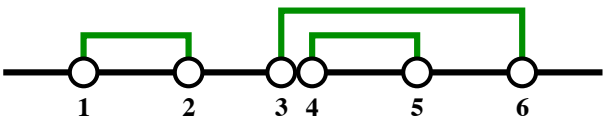
$$\rho(E) \;\to\; i\,\frac{(\pm)^{N_p}}{S}\,\Big(\prod\gamma\Big)\; R^{(0)}(E)\,G^{(0)}\,R^{(0)}(E+X_{M_1})\,\ldots\,G^{(0)}\,R^{(0)}(E+X_{M_k})\,\rho_{t=0} \quad , \tag{63}$$

where $\prod\gamma$ denotes the product over all contractions (59), N_p is the number of permutations of reservoir field superoperators to disentangle the fermionic contractions, and $S = \prod_i m_i!$ is a symmetry factor arising for the case when pairs of vertices are connected by m_i contractions.

The determination of the shift variables X_{M_i} is simplified by noting that, according to (59), a single contraction $\gamma_{12}^{p_1p_2}$ between $A_1^{p_1}$ and $A_2^{p_2}$ is only possible for $\eta_1 = -\eta_2$, $\alpha_1 = \alpha_2$ and $\omega_1 = \omega_2$. This gives $\bar{\omega}_{12} = \bar{\mu}_{12} = 0$, i.e. if the two indices fall both into the same set M_i, there is no contribution to the shift X_{M_i}. As a consequence, the left index 1 of a contraction $\gamma_{12}^{p_1p_2}$ will contribute only to those resolvents, which stand between the two field operators $A_1^{p_1}$ and $A_2^{p_2}$ in the original series. For this reason, the last resolvent in (63) has no shift $X_{M_k} = 0$ and is given by $R^{(0)}(E)$.

With the diagrammatic rules it is straightforward to translate a particular diagram, which we visualize as follows:

$\rho(E) \;\to\; i$



$\rho_{t=0}$

$$\begin{aligned} =\; i\,\gamma_{12}^{p_1p_2}\,\gamma_{36}^{p_3p_6}\,\gamma_{45}^{p_4p_5}\,R^{(0)}(E)\,G_1^{(0)p_1}\,R^{(0)}(E_1+\bar{\omega}_1)\,G_2^{(0)p_2}\,R^{(0)}(E) \\ G_{34}^{(0)p_3p_4}\,R^{(0)}(E_{34}+\bar{\omega}_{34})\,G_5^{(0)p_5}\,R^{(0)}(E_3+\bar{\omega}_3)\,G_6^{(0)p_6}\,R^{(0)}(E)\,\rho_{t=0} \quad , \end{aligned} \tag{64}$$

where we used the notation $E_{1\ldots n} = E + \bar{\mu}_{1\ldots n}$, see (38). In the diagrams, the green lines are the contractions, the circles denote the vertices, and the black lines connecting the vertices represent the resolvents $R^{(0)}$ describing the dot propagation in Fourier space. The indices of the shift

variables of a particular resolvent can be determined by drawing a vertical line at the position of that resolvent and taking the left indices of all contractions which cut through this line. We note that we do not distinguish between diagrams which differ only by a permutation of the contractions connected to a certain vertex, i.e. the permutation of the two green lines connected to the indices 3 and 4 in the above example does not lead to a new diagram.

To bring the density matrix $\rho(E)$ into the form (27) and to identify the effective Liouvillian $L(E)$, we note that each diagram consists of a sequence of connected blocks, defined by the property that each vertical line will at least hit one contraction, connected by resolvents $R^{(0)}(E)$. E.g., the diagram (64) consists of a sequence of two blocks. Denoting the sum of all connected diagrams by $\Sigma(E)$, the diagrammatic series can be written as

$$\rho(E) \;=\; i\,\big\{R^{(0)}(E) \;+\; R^{(0)}(E)\,\Sigma(E)\,R^{(0)}(E) + \\ + R^{(0)}(E)\,\Sigma(E)\,R^{(0)}(E)\,\Sigma(E)\,R^{(0)}(E) \;+\; \ldots\big\}\;\rho_{t=0} \;=\; \frac{i}{E \;-\; L^{(0)} \;-\; \Sigma(E)}\,\rho_{t=0} \quad . \tag{65}$$

Comparing to (27), we see that the effective Liouvillian is given by

$$L(E) \;=\; L^{(0)} \;+\; \Sigma(E) \quad , \tag{66}$$

and $\Sigma(E)$ consists of the sum of all connected diagrams with translation rule

$$\Sigma(E) \;\to\; \frac{(\pm)^{N_p}}{S}\,\Big(\prod\gamma\Big)_{con}\,G^{(0)}\,R^{(0)}(E_{M_1}+\bar\omega_{M_1})\,\ldots\,G^{(0)}\,R^{(0)}(E_{M_k}+\bar\omega_{M_k})\,G^{(0)} \quad , \tag{67}$$

where $(\prod\gamma)_{con}$ means that only connected diagrams are considered. E.g. some of the lowest order diagrams of $\Sigma(E)$ are given by

$\Sigma(E) \;=$  $+\ \ldots$ (68)

$\Sigma(E)$ is the dissipative part of the effective Liouvillian, which contains the whole information of the coupling to the reservoirs and leads to irreversible time evolution. In time space we obtain $L(t) = L^{(0)}\delta(t-0^+) + \Sigma(t)$, such that the kinetic equation (23) reads

$$i\dot\rho(t) \;=\; L^{(0)}\,\rho(t) \;+\; \int_0^t dt'\,\Sigma(t-t')\,\rho(t') \quad . \tag{69}$$

The first term describes the von Neumann equation in the absence of the reservoirs, whereas the second one is the dissipative part. The two terms are the analog of the "flow" and the "collision" term of quantum Boltzmann equations.

Local observables. From the density matrix $\rho(E)$ the time evolution of all averages of local observables can be calculated. The diagrammatic expansion can also be formulated for the calculation of arbitrary observables containing reservoir degrees of freedom or correlation functions. E.g., if an observable I of the generic form (10) is taken

$$I \;=\; \frac{1}{n!}\left\{\begin{array}{c} 1 \\ \eta_1\eta_2\ldots\eta_n \end{array}\right\} :a_n a_{n-1}\ldots a_1: \; i_{12\ldots n} \;\to\; \frac{1}{n!}\,i_{12\ldots n}\;:a_1 a_2\ldots a_n: \quad , \tag{70}$$

we define a corresponding superoperator L_I by the anticommutator

$$L_I \;=\; \frac{i}{2}\,[I,\cdot]_+ \;=\; \frac{1}{n!}\sum_{p=\pm} I^{(0)p\ldots p}_{1\ldots n}\;:A^p_1\ldots A^p_n:\quad , \tag{71}$$

with

$$I^{(0)p\ldots p}_{1\ldots n}\,b \;=\; \frac{i}{2}\left\{\begin{array}{ll} 1 & \text{for } n \text{ even} \\ \sigma^p & \text{for } n \text{ odd}\end{array}\right\}\left\{\begin{array}{ll} i_{1\ldots n}\,b & \text{for } p=+ \\ b\,i_{1\ldots n} & \text{for } p=-\end{array}\right\}\quad , \tag{72}$$

such that the average can be written as

$$\langle\, I\,\rangle(t) \;=\; \mathrm{Tr}_{\rm tot}\, I\,\rho_{\rm tot}(t) \;=\; -i\,\mathrm{Tr}\,\mathrm{Tr}_{\rm res}\, L_I\, e^{-iL_{\rm tot}t}\,\rho_{t=0}\,\rho^{\rm eq}_{\rm res}\quad . \tag{73}$$

This expression has a formal similiarity to

$$\begin{aligned}\dot\rho(t) \;&=\; -i\,\mathrm{Tr}_{\rm res}\, L_{\rm tot}\, e^{-iL_{\rm tot}t}\,\rho_{\rm tot}(t=0) \;=\; -i\,\mathrm{Tr}_{\rm res}(L^{(0)}+L_{\rm res}+L_V)\, e^{-iL_{\rm tot}t}\,\rho_{t=0}\,\rho^{\rm eq}_{\rm res} \\ &=\; -i\,L^{(0)}\,\rho(t) \;-\; i\,\mathrm{Tr}_{\rm res} L_V\, e^{-iL_{\rm tot}t}\,\rho_{t=0}\,\rho^{\rm eq}_{\rm res}\quad ,\end{aligned} \tag{74}$$

where $\mathrm{Tr}_{\rm res}L_{\rm res}=0$ has been used in the last line. Comparing to the kinetic equation (69), we find

$$-\,i\,\mathrm{Tr}_{\rm res} L_V\, e^{-iL_{\rm tot}t}\,\rho_{t=0}\,\rho^{\rm eq}_{\rm res} \;=\; -i\int_0^t dt'\,\Sigma(t-t')\,\rho(t')\quad . \tag{75}$$

Therefore, when applying the same perturbative expansion to (73), we obtain the result

$$\langle\, I\,\rangle(t) \;=\; -i\int_0^t dt'\,\mathrm{Tr}\,\Sigma_I(t-t')\,\rho(t')\quad ,\quad \langle\, I\,\rangle(E) \;=\; -i\,\mathrm{Tr}\,\Sigma_I(E)\,\rho(E)\quad , \tag{76}$$

with the only difference that the first vertex of the kernel $\Sigma_I(E)$ has to be the vertex $I^{(0)}$ instead of $G^{(0)}$, i.e. the diagrammatic rule (67) changes to

$$\begin{aligned}\Sigma_I(E) \;\to\; &\frac{(\pm)^{N_p}}{S}\,\Big(\prod\gamma\Big)_{con} \\ &I^{(0)}\,R^{(0)}(E_{M_1}+\bar\omega_{M_1})\,G^{(0)}\,R^{(0)}(E_{M_2}+\bar\omega_{M_2})\,\ldots\,G^{(0)}\,R^{(0)}(E_{M_k}+\bar\omega_{M_k})\,G^{(0)}\quad .\end{aligned} \tag{77}$$

A prominent example for an observable is the particle current operator flowing from reservoir α into the local system defined by

$$I_\alpha \;=\; -\frac{d}{dt}\,N_\alpha \;=\; -i\,[H_{\rm tot},N_\alpha] \;=\; -i\,[V,N_\alpha]\quad . \tag{78}$$

Inserting the form (10) of V, one finds after some straightforward manipulations

$$(i_\alpha)_{1\ldots n} \;=\; i\sum_{k=1}^n \eta_k\,\delta_{\alpha_k\alpha}\,g_{1\ldots n}\quad ,\quad (I_\alpha)^{(0)p\ldots p}_{1\ldots n} \;=\; -\frac{1}{2}\sum_{k=1}^n \eta_k\,\delta_{\alpha_k\alpha}\,p\,G^{(0)p\ldots p}_{1\ldots n}\quad . \tag{79}$$

Using the diagrammatic expansion (67) and (77), the kernels $\Sigma(E)$ and $\Sigma_I(E)$ can be calculated in perturbation theory w.r.t. the bare vertices $G^{(0)}$, and the local density matrix $\rho(E)$ and the average $\langle I\rangle(E)$ of any observable I follow from (65) and (76) in Fourier space. Using inverse

Fourier transform the time evolution can finally be discussed following the strategy described in Section 3. Stationary quantities are obtained from

$$\rho_{\rm st} \;=\; \lim_{t\to\infty}\,\rho(t) \;=\; \lim_{E\to 0^+}\,(-iE)\,\rho(E) \quad\Leftrightarrow\quad L(E=0^+)\,\rho_{\rm st} \;=\; 0 \tag{80}$$

$$\langle I\rangle_{\rm st} \;=\; \lim_{t\to\infty}\,\langle I\rangle(t) \;=\; \lim_{E\to 0^+}\,(-iE)\,\langle I\rangle(E) \;=\; -i\,{\rm Tr}\,\Sigma_I(E=0^+)\,\rho_{\rm st} \quad . \tag{81}$$

Applications of these perturbative schemes for the calculation of transport properties will be discussed in the lecture C7 by M. Wegewijs. Similiar schemes have also been developed to calculate correlation functions [21] and to consider explicitly time-dependent Hamiltonians [9, 22]. Concerning the latter first applications have considered adiabatic response [22] and quantum quenches [9] for the IRLM.

Analytic properties. From the perturbative expansion one finds that the effective Liouvillian $L(E)$ has a branch cut on the real axis and is analytic in the upper and lower half of the complex plane. This can be seen from the resolvents $R^{(0)}(E_M+\bar\omega_M)$ since $L^{(0)}=[H^{(0)},\cdot]$ is a self-adjoint superoperator with real eigenvalues and all frequency variables $\bar\omega$ are integrated over the real axis. The same analytic property holds for the resolvent $R(E)=\frac{1}{E-L(E)}$, since, due to (27), we get for any initial density matrix $\rho_{t=0}$

$$\begin{aligned} R(E)\,\rho_{t=0} \;&=\; -i\,\rho(E) \;=\; -i\int_0^\infty dt\,e^{iEt}\,\rho(t) \;=\; -i\int_0^\infty dt\,e^{iEt}\,{\rm Tr}_{\rm res}\,\rho_{\rm tot}(t) \\ &=\; -i\int_0^\infty dt\,e^{iEt}\,{\rm Tr}_{\rm res}\,e^{-iL_{\rm tot}t}\,\rho_{\rm tot}(t=0) \;=\; {\rm Tr}_{\rm res}\,\frac{1}{E\,-\,L_{\rm tot}}\,\rho_{\rm tot}(t=0) \quad . \end{aligned} \tag{82}$$

This function can only have a branch cut on the real axis since $L_{\rm tot}=[H_{\rm tot},\cdot]$ is a self-adjoint superoperator with real eigenvalues. To calculate the time evolution we have seen from (28) that the integration $\int dE$ is slightly above the real axis and has to be closed in the lower half of the complex plane (due to $t>0$). It is very inconvenient to calculate this integral by enclosing the branch cut of the integrand on the real axis due to the rapidly oscillating function e^{-iEt} in the integrand on the scale $1/t$. Therefore, in analogy to the standard procedure for response functions, one tries to find an appropriate analytic continuation of the functions $L(E)$ and $R(E)$ into the lower half of the complex plane such that all branch cuts point into the direction of the negative imaginary axis starting at certain singularities z_n with ${\rm Im}z_n\le 0$. We achieve this in two steps. First, we will transform the perturbative series for $L(E)$ into a self-consistent equation by resumming all blocks of connected diagrams on the propagators connecting the vertices. The diagrammatic representation allows this to be done in a unique way and, as a result, the bare resolvents $R^{(0)}(E_M+\bar\omega_M)$ are replaced by the full ones $R(E_M+\bar\omega_M)$ and no diagrams are allowed with connected sub-blocks without any free lines on the propagators which we indicate by $(\prod\gamma)_{irr}$

$$\Sigma(E) \;\to\; \frac{(\pm)^{N_p}}{S}\,\Big(\prod\gamma\Big)_{irr}\,G^{(0)}\,R(E_{M_1}+\bar\omega_{M_1})\,\ldots\,G^{(0)}\,R(E_{M_k}+\bar\omega_{M_k})\,G^{(0)} \quad . \tag{83}$$

For ${\rm Im}E>0$, all resolvents are analytic functions w.r.t. the integration variables $\bar\omega_i$ in the upper half of the complex plane. Therefore, in the second step, we can close all integration contours in the upper half and have to enclose only the nonanalytic features arising from the spectral function $\rho_{\alpha\sigma}(\omega)$ and the Bose/Fermi distribution $f_\alpha(p'\eta\omega)$ in the contraction $\gamma^{pp'}_{11'}$, defined in (59). Thereby we assume that the frequency dependence of the vertices $G^{(0)p\ldots p}_{1\ldots n}$ can be neglected.

Decomposing the Bose/Fermi distribution in symmetric and antisymmetric parts and using the representation in terms of the Masubara frequencies $\omega_n^\alpha = 2n\pi T_\alpha$ ($\omega_n^\alpha = (2n+1)\pi T_\alpha$) for bosons (fermions), we can write the contraction in the form

$$\gamma_{11'}^{pp'} = \delta_{1\bar{1}'}\, p' \left\{ \begin{array}{c} \eta \\ 1 \end{array} \right\} \bar{\rho}_{\alpha\sigma}(\bar{\omega}) \left\{ \mp \frac{1}{2} + p' \left(f_\alpha(\bar{\omega}) \pm \frac{1}{2} \right) \right\} = \delta_{1\bar{1}'} \left(p'\, \gamma_1^s + \gamma_1^a \right) \quad , \tag{84}$$

$$\gamma_1^s = \mp \frac{1}{2} \left\{ \begin{array}{c} \eta \\ 1 \end{array} \right\} \bar{\rho}_{\alpha\sigma}(\bar{\omega}) \quad , \tag{85}$$

$$\begin{aligned} \gamma_1^a &= \left\{ \begin{array}{c} \eta \\ 1 \end{array} \right\} \bar{\rho}_{\alpha\sigma}(\bar{\omega}) \left\{ f_\alpha(\bar{\omega}) \pm \frac{1}{2} \right\} \\ &= \left\{ \begin{array}{c} \eta \\ 1 \end{array} \right\} \bar{\rho}_{\alpha\sigma}(\bar{\omega})\, T_\alpha \frac{1}{2} \sum_n \left(\frac{1}{\bar{\omega} - i\omega_n^\alpha} + \frac{1}{\bar{\omega} + i\omega_n^\alpha} \right) \quad , \end{aligned} \tag{86}$$

where $\bar{\rho}_{\alpha\sigma}(\bar{\omega}) = \rho_{\alpha\sigma}(\omega)$. Thus, after performing all integrations $\int d\bar{\omega}_i$, and assuming for the moment that the spectral function $\bar{\rho}_{\alpha\sigma}(\bar{\omega})$ is an analytic function in the upper half, the quantities $\bar{\omega}_M$ occuring in the resolvents $R(E_M + \bar{\omega}_M)$ will consist of a sum of positive Matsubara frequencies $i \sum_{j\in M} |\omega_{n_j}^{\alpha_j}|$. As a consequence, the analytic continuation w.r.t. E of this result for $L(E)$ into the lower half of the complex plane will lead to nonanalytic features at $E_M + i \sum_{j\in M} |\omega_{n_j}^{\alpha_j}| = z_k^p$, where z_k^p are the poles of the resolvent $R(E)$ after the analytic continuation into the lower half of the complex plane. Since $R(E)$ is analytic in the upper half, the poles z_k^p have to lie in the lower half, and we find that $L(E)$ has an infinite series of poles in the lower half located at

$$E = z_k^p - \bar{\mu}_M - i \sum_{j\in M} |\omega_{n_j}^{\alpha_j}| \quad , \tag{87}$$

which, at zero temperature, turn into a series of branch cuts in the direction of the negative imaginary axis with branching points z_n located at the poles z_k^p shifted by any combinations of the chemical potentials of the reservoirs

$$z_n = z_k^p - \bar{\mu}_M \quad . \tag{88}$$

This result has formed the basis for the generic discussion of the time evolution in Section 3.

Influence of spectral function. We note that the spectral function $\bar{\rho}_{\alpha\sigma}(\bar{\omega})$ of the models introduced in Section 2 does not change this picture. For quantum dots coupled to Fermi liquid leads, like the Kondo model or the IRLM, the spectral function $\bar{\rho}_{\alpha\sigma}(\bar{\omega}) = \frac{D^2}{D^2+\bar{\omega}^2}$ defines just a high-energy cutoff function with pole at $\bar{\omega} = iD$ in the upper half and residuum $-iD/2$. The contribution of this pole to the frequency integration leads for $D \to \infty$ either to a vanishing or to a regular contribution in E. For the ohmic spin boson model we get from (19) that $\bar{\rho}(\bar{\omega}) = 2\alpha|\bar{\omega}|\theta(\eta\bar{\omega})\frac{D^2}{D^2+\bar{\omega}^2}$, which has a branch cut on the whole imaginary axis. However, since the vertex $g_1 = \frac{1}{2}\sigma_z$ is independent of η, we can sum the contraction $\gamma_{11'}^{pp'}$ over η and η' at fixed $\bar{\omega}$ and $\bar{\omega}'$ (which are the integration variables) and get from (59) for the case of bosons the effective contraction

$$\gamma_{11'}^{pp'} = \delta(\bar{\omega} + \bar{\omega}')\, p'\, 2\, \alpha\, \bar{\omega}\, \frac{D^2}{D^2 + \bar{\omega}^2}\, f(p'\bar{\omega}) = \delta(\bar{\omega} + \bar{\omega}') \left(p'\gamma_1^s + \gamma_1^a \right) \quad , \tag{89}$$

with

$$\gamma_1^s = -\alpha\, \bar{\omega}\, \frac{D^2}{D^2 + \bar{\omega}^2} \quad , \quad \gamma_1^s = \alpha\, \bar{\omega}\, \frac{D^2}{D^2 + \bar{\omega}^2} \left(2f(\bar{\omega}) + 1 \right) \quad , \tag{90}$$

where the index $1 \equiv \bar{\omega}$ contains only the frequency variable. Since $\bar{\omega}$ is an analytic function, there is no change of the analytic structure of $L(E)$ for the ohmic spin boson model. For a generic frequency dependence of the spectral function, the analytic structure might change. If $\bar{\rho}(\bar{\omega})$ has a branch cut in the upper half in the direction of the positive imaginary axis starting at $\Delta_\rho + i\gamma_\rho$, with $\gamma_\rho \geq 0$, the position (88) of the branching points of $L(E)$ can be shifted by multiples of $-\Delta_\rho - i\gamma_\rho$. This can e.g. happen for superconducting leads, where Δ_ρ corresponds to the superconducting gap and $\gamma_\rho = 0$. For sub- or super-ohmic spin boson models there is no change of the analytic properties since the branch cuts of the spectral function start at the origin.

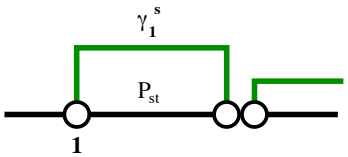
Symmetric part of the contraction. We note that the part γ_1^s of the contraction (84) involving the symmetric part of the Bose/Fermi distribution plays a special role. It is the only part of the contraction $\gamma_{11'}^{pp'}$ which depends on the Keldysh indices via p' and it depends on the frequency only via the spectral function. In particular for a spectral function of the form $\bar{\rho}(\bar{\omega}_1) = \frac{D^2}{D^2+\bar{\omega}_1^2}$, i.e. if it just acts as a high-energy cutoff function but has no other special form, the frequency integration $\int d\bar{\omega}_1$ will involve only the pole of the spectral function at $\bar{\omega}_1 = iD$ when closed in the upper half. In the limit $D \to \infty$ this means that this integration gives either zero (if more than one resolvent involves $\bar{\omega}_1$) or a constant if this contraction connects two consecutive vertices

$$\int d\bar{\omega}_1 \, \frac{D^2}{D^2+\bar{\omega}_1^2} \, R(E_{1\ldots n} + \bar{\omega}_{1\ldots n}) \; = \; \pi \, D \, R(E_{1\ldots n} + \bar{\omega}_{2\ldots n} + iD) \; \xrightarrow{D\to\infty} \; -i\,\pi \quad . \tag{91}$$

As a result, the symmetric part of the contraction can be integrated out analytically and can be incorporated in an effective vertex by taking the two consecutive vertices together to a single one. The same can be shown for the ohmic spin boson model [10] due to its special algebra, whereas for more general spectral functions with nonanalytic features in the upper half this is not the case. If it holds, one important consequence of this property is that the special pole at $E_M + \bar{\omega}_M = z_{\rm st}^p = 0$ of the resolvents $R(E_M + \bar{\omega}_M)$ leads to regular contributions in the limit $D \to \infty$ and does not contribute to the branch cuts of $L(E)$. The reason is the special form $P_{\rm st}(E) = |x_{\rm st}(E)\rangle$ Tr for the projector of the mode $k = st$ (see Eq. (34)) together with the property

$$\sum_p \, {\rm Tr}\, G_{1\ldots n}^{(0)p\ldots p} \; = \; 0, \tag{92}$$

which follows straightforwardly from the definition (55). Since the contractions $\gamma_{11'}^{pp'}$ are independent of the first Keldysh index p (see Eq. (59)), this means that if the projector $P_{\rm st}(E_M + \bar{\omega}_M)$ is inserted between two consecutive vertices, at least one of the contractions $\gamma_{11'}^{pp'}$ associated with the right vertex must point into the left direction and only its p'-dependent symmetric part $p'\gamma_1^s$ will contribute



(93)

As shown above this means that this contraction has to connect the two consecutive vertices and the frequency integration gives a constant. Thus, the case $k = $ st does not contribute to the positions (88) of branch cuts for $L(E)$ or $R(E)$. There might be other accidental poles $z_k^p = 0$ for $k \neq st$, like e.g. for multi-channel Kondo models with non-Fermi liquid behaviour [6], but

for the models discussed in Section 2, this is not the case. Therefore, for these models, the pole at $z_{\rm st} = 0$ is isolated, as already stated in Section 3 after Eq. (45).

Breakdown of perturbation theory for time-evolution problems. Finally we note that it is very important to use the perturbative expansion of $L(E)$ in the self-consistent form (83) in order to find the right position of the branching points z_n of the branch cuts of $R(E)$. For the original series (67) involving the bare resolvent $R^{(0)}(E_M + \bar{\omega}_M)$, the same considerations as above lead to branch cuts of $L(E)$ starting on the real axis at the value $z_n = \lambda^{(0)} - \bar{\mu}_M$, where $\lambda^{(0)}$ is a real eigenvalue of the bare Liouvillian $L^{(0)}$. This would have a dramatic effect on the long-time evolution because it leads to non-exponential decay. However, this result is not correct since perturbation theory is very dangerous in the regime $|E - z_n| \sim \Gamma$, where Γ is a typical decay rate. At low frequencies, the resolvents can then become very large of the order of the inverse coupling constant, raising serious questions about convergence. In particular for the original series (67) involving $R^{(0)}$ this effect is most dramatic since a series of connected sub-blocks contains an arbitrary number of resolvents with exactly the same argument, i.e. the singularity at low frequency appears to an arbitrary power. Such a series is certainly not convergent and it is necessary to resum it first to the self-consistent version (83) before determining the position of the branching points. E.g. consider a contribution to the effective Liouvillian of the form (α is some small dimensionless coupling constant and Δ denotes a typical low-energy scale)

$$\alpha\,(-iE + \alpha\Delta)\,\ln\frac{D}{-iE + \alpha\Delta} = \alpha\,(-iE + \alpha\Delta)\,\ln\frac{D}{-iE} - \alpha^2\,\Delta - \frac{1}{2}\,i\,\alpha^3\,\frac{\Delta^2}{E} + O(\alpha^4) \quad . \tag{94}$$

The logarithm on the l.h.s. has a branching point at $E = -i\alpha\Delta$ but the expanded form gives in $O(\alpha)$ and $O(\alpha^2)$ a branching point at $E = 0$. The mistake can only be seen by considering higher orders in α, where an infinite series of terms with a pole at $E = 0$ is obtained. Due to the factor $-iE + \alpha\Delta$ in front of the logarithm, this artifact is even not visible in $O(\alpha^2)$ but starts in $O(\alpha^3)$ or higher. The form (94) arises e.g. for the Kondo model and for the ohmic spin boson model, where we will show in Section 5 that $\frac{\partial}{\partial E}L(E)$ must be a slowly varying logarithmic function leading to typical terms of the form (94), see also Eq. (105). For the ohmic spin boson model previous calculations [2] have predicted terms with non-exponential decay which have been corrected recently [10, 23, 24].

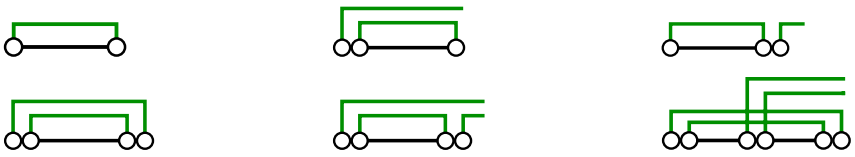
5 Renormalization group

General remarks. At low temperatures the perturbative calculation of the effective Liouvillian can break down even at small reservoir-system coupling for two reasons. First, at high energies (the so-called ultraviolett regime), the frequency integrals are typically logarithmic leading to logarithmic contributions $\sim \alpha^k(\ln\frac{D}{E-z_n})^l$ in higher-order perturbation theory, with $k \geq l$, where z_n are the branching points (88) of the resolvent $R(E)$ and α is an appropriate dimensionless coupling constant. Secondly, even if perturbation theory does not contain ultraviolett logarithmic divergencies in the limit $D \to \infty$, it may contain logarithmic terms $\sim \alpha \ln\frac{E-z_n}{E-z_m}$, which, for $E \to z_n$, turn into the form $\sim \alpha \ln\frac{E-z_n}{z_n-z_m}$, which can lead to a breakdown of perturbation theory at low energies (the so-called infrared regime). Therefore, a method is needed capable of reorganizing perturbation theory such that all ultravioltett and infrared logarithmic divergencies are resummed. Concerning high energies, resumming all logarithmic contributions

$\sim \alpha^k (\ln \frac{D}{E-z_n})^l$ with $l = k, k-1, k-2, \ldots$ is called leading order, sub-leading order, sub-sub leading order, etc. approximation (sometimes also referred to as 1-loop, 2-loop, 3-loop, etc.). In traditional (so-called poor man scaling) RG methods [25], one tries to perform this resummation by integrating out high-energy scales, i.e. the band width D is successively reduced in infinitesimal steps and the physical quantity of interest is kept invariant by renormalizing the coupling constants and other energy scales. Provided that the renormalized coupling constants remain small (the so-called weak-coupling regime), a well-controlled truncation scheme can be set up by neglecting higher-order terms in the renormalized couplings. This strategy has also been used for calculating stationary quantities of nonequilibrium problems [18], but it turns out that for the calculation of the effective Liouvillian $L(E)$ it is very hard to set up a systematic truncation scheme. This has been improved by using a high-energy cutoff on the imaginary axis by cutting off the Matsubara frequencies of the Bose/Fermi distribution function [3] and applied to various models [3, 5, 7, 19, 21]. Here we will follow another route by describing the E-RTRG method [6], which is unique in the sense that it is capable of dealing with all logarithmic divergencies at high *and* low energies. Technically, this is achieved by considering the perturbation theory not for $L(E)$ but for its first or second derivative w.r.t. the Fourier variable E together with a proper resummation in terms of effective vertices. Whether a first or a second derivative is needed depends on the model under consideration. This leads to a series where all frequency integrals converge at high energies and the limit $D \to \infty$ can be performed in all orders. As a consequence one obtains a *universal* differential equation (called RG equation) for $L(E)$ independent of the specific choice of the high-energy cutoff function. Furthermore, the RG equation for $L(E)$ turns out to be such that the divergence at low energies for $E \to z_n$ is at most $\sim \frac{1}{E-z_n}$ multiplied with a perturbative series in terms of effective vertices which exists in the limit $E \to z_n$. This allows for a systematic solution at low energies as well. Besides the effective Liouvillian also effective vertices will appear in the RG equation due to the resummation procedure, for which similiar universal RG equations can be derived. Provided that the effective vertices stay small (so-called weak coupling problems) the RG equations can be systematically truncated and well-controlled universal properties can be determined at high as well as at low energies. According to the discussion in Section 3 this allows a well-controlled discussion of the time evolution at short and long times together with the crossover behaviour. The high-energy cutoff D will only appear in the initial condition for the various quantities which are calculated by a well-controlled perturbation theory in the bare couplings at $E = iD$. This procedure has the advantage that by construction only the universal properties of the model are obtained, although it is also possible to keep D fixed and solve the RG equations for a given high-energy cutoff function. Furthermore, the use of a physical scale E as flow parameter of the RG equations has the advantage that at each stage of the flow the solution $L(E)$ provides a result for a physical quantity. Moreover, since E is a complex flow parameter, the flow can be solved on any path in the complex plane which is very helpful to find appropriate analytic continuations of retarded functions into the lower half of the complex plane, even by using numerical methods, see also the discussion in Section 3 after Eq. (38).

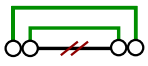
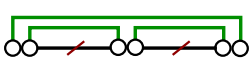
Derivation of the E-RTRG equations. To illustrate the general strategy for the derivation of the RG equations within the E-RTRG method we consider here, for simplicity, a spectral function of the form $\bar{\rho}(\bar{\omega}) = \frac{D^2}{D^2+\bar{\omega}^2}$, which arises typically for fermionic metallic reservoirs where the d.o.s. is approximately a constant in the physically relevant energy regime. Therefore, we consider only the fermionic case in the following. Furthermore, we assume that only 1- and 2-point vertices occur in the original model and that the frequency dependence of the bare

vertices g_1 and g_{12} can be neglected. This applies to the Kondo model and the IRLM introduced in Section 2. For the ohmic spin boson model, a similiar procedure can be used to derive the RG equations, see Ref. [10]. For such models, one obtains a problem with convergence at high energies if the number of frequency integrations is larger or equal to the number of resolvents where the frequencies occur. For models with 1- and 2-point vertices this means that diagrammatic sub-elements of the form



(95)

lead to problems at high energies and have to be avoided. This can be achieved by taking a single or a double derivative w.r.t. E of the resolvents occuring in these diagrams. Therefore, the idea is to consider a perturbative expansion for the derivatives $\frac{\partial}{\partial E}L(E)$ or $\frac{\partial^2}{\partial E^2}L(E)$ and to resum the series such that no subelements of the form (95) remain. The procedure is quite straightforward and we illustrate it for the case of a model where only 2-point vertices occur, like e.g. the Kondo model. Here, to guarantee convergence we consider two derivatives w.r.t. E of the diagrammatic series (83) of $L(E)$ in the self-consistent form. Since the E-dependence occurs only in the resovents $R(E_M + \bar{\omega}_M)$, we can either take two derivatives of a single resolvent or two single derivatives of different resolvents. Fixing the positions of the resolvents we can then resum all remaining diagrams in a unique way such that the bare 2-point vertices $G_{12}^{(0)pp}$ are replaced by full effective 2-point vertices $G_{12}^{p_1p_2}(E)$, which are defined as the sum of all connected diagrams with 2 external reservoir lines. With the convention that these two external lines are directed to the right, it turns out that the energy argument of an effective vertex is identical to the one of the preceding resolvent, i.e. only the combination $R(E_M + \bar{\omega}_M)G_{12}^{p_1p_2}(E_M + \bar{\omega}_M)$ can occur in the diagrammatic expansion. Furthermore, for all diagrams contributing to the effective vertex $G_{12}^{p_1p_2}(E)$, where the two external lines have the sequence 21, a fermionic sign has to be added. After this resummation, the diagrammatic series for $\frac{\partial^2}{\partial E^2}L(E)$ up to third order in the effective vertices reads

$$\frac{1}{2}\frac{\partial^2}{\partial E^2}L(E) = \frac{1}{2}\frac{1}{2}$$  $$+$$  $$+ \mathcal{O}\left(G^4\right) \quad , \tag{96}$$

where the red slash indicates a derivative $\frac{\partial}{\partial E}$ of the corresponding resolvent (two slashes indicate the second derivative $\frac{\partial^2}{\partial E^2}$). This is one of the central equations in the E-RTRG approach. Prefactors arsing from the symmetry factor $\frac{1}{S}$ have explicitly been indicated and all vertices are full effective 2-point vertices from now on. All frequency integrations are convergent even if one neglects the frequency-dependence of the effective vertices (those can only enhance convergence). Therefore on the r.h.s. of this differential equation we can take the limit $D \to \infty$. This property holds in all orders since, by construction, all diagrammatic subelements (95) leading to a divergence in the infinite-D limit have been eliminated by the resummation procedure. To close the equation one can also derive in the same way a differential equation for the effective

2-point vertex

$$\frac{\partial}{\partial E}G_{12}^{p_1p_2}(E) \;=\; \Big[\;\; - (1\leftrightarrow 2)\Big]$$

$$+\;\frac{1}{2}\;\;+\;\frac{1}{2}\;\;$$

$$+\;\Big[\;\;+\;\;- (1\leftrightarrow 2)\Big] + O(G^4)\;. \qquad (97)$$

After the limit $D\to\infty$ has been taken, the symmetric part (85) of the contraction becomes an analytic function and does not contribute to the frequency integration when closing the integration in the upper half of the complex plane. This means that the Keldysh indices no longer appear explicitly in the RG equations, i.e. only the effective 2-point vertices averaged over the Keldysh indices are needed, which we denote by $G_{12}(E)=\sum_{p_1p_2}G_{12}^{p_1p_2}(E)$. This simplifies the analysis considerably. As a result all contractions can be replaced by the antisymmetric part given by (86)

$$\gamma_{11'}^{pp'}\;\to\;\gamma_1^a\;=\;f_\alpha(\bar\omega)\;-\;\frac{1}{2}\quad, \qquad (98)$$

where we have already taken the limit $D\to\infty$ and integrated out the trivial part $\delta_{1\bar 1}$ of all contractions in the RG diagrams. By convention, $\bar\omega$ is always the frequency variable of the left vertex.

Frequency dependence. To calculate the integrals over the internal frequencies in the RG diagrams, it is necessary to know the frequency dependence of the effective vertices and the Liouvillian. This can be treated systematically by the formalism. Provided that the bare vertices are frequency-independent, one finds for the vertices that the diagrammatic series for the difference $G_{12}(E)-G_{12}(E)_{\bar\omega_1=\bar\omega_2=0}$ can be resummed by a similiar procedure in terms of effective 2-point vertices such that the limit $D\to\infty$ is well-defined. The reason is that at least one resolvent in the original perturbative series must involve the difference $R(E_M+\bar\omega_M+\bar\omega_{M_{\rm ex}})-R(E_M+\bar\omega_M)$, where $M_{\rm ex}$ contains some of the external indices $\{1,2\}$. Fixing this resolvent and resumming the rest of the diagram in terms of effective 2-point vertices yields in lowest order the equation

$$\;=\;+\;-\;+\;\mathcal{O}\left(G^3\right)\quad, \qquad (99)$$

where the filled double dots represent the effective vertices at zero frequency. This is the second key equation in the E-RTRG approach. A contraction with an open circle and external frequency $\bar\omega_i$ indicates that the resolvent corresponding to the vertical cut at the position of that circle has to be replaced by the difference $R(E_M+\bar\omega_M+\bar\omega_i)-R(E_M+\bar\omega_M)$. This difference falls off $\sim(\bar\omega_M)^2$ w.r.t. the internal frequency integration variables $\bar\omega_M$ and, therefore, all frequency integrations are convergent in the limit $D\to\infty$. For the frequency dependence of the Liouvillian $L(E_M+\bar\omega)$ it turns out that the similiar diagrammatic series for the difference $L(E+\bar\omega)-L(E)$ does not exist in the limit $D\to\infty$, similiar to the fact that two derivatives are needed for convergence (see above). Therefore, one defines a discrete version of the second derivative $\Delta^2_{\bar\omega}L(E)$ via

$$L(E+\bar\omega)\;=\;L(E)\;+\;\frac{\partial}{\partial E}L(E)\,\bar\omega\;+\;\Delta^2_{\bar\omega}L(E)\;=\;L(E)\;+\;\frac{\partial}{\partial E}L(E)\,\bar\omega\;+\;O(G^2)\quad, \qquad (100)$$

and finds that $\Delta_{\bar{\omega}}^2 L(E)$ exists in the limit $D \to \infty$ and is at least of $O(G^2)$ since it involves second and higher-order derivatives of the Liouvillian. Neglecting $O(G^2)$ (note that this contributes $O(G^4)$ to the RG equations (96) and (97)), the resolvents occuring in the RG diagrams of (96) and (97) can be written as

$$R(E_M + \bar{\omega}_M) \;=\; \frac{1}{\bar{\omega}_M \,+\, \chi(E_M)}\, Z(E_M) \,+\, O(G^2) \quad , \tag{101}$$

with

$$\chi(E) \;=\; Z(E)\,(E \,-\, L(E)) \quad , \quad Z(E) \;=\; \frac{1}{1 \,-\, \frac{\partial}{\partial E} L(E)} \quad , \tag{102}$$

where the RG equation for $Z(E)$ follows from the one for $\frac{\partial^2}{\partial E^2} L(E)$ by

$$\frac{\partial}{\partial E} Z(E) \;=\; Z(E) \left\{ \frac{\partial^2}{\partial E^2} L(E) \right\} Z(E) \;\sim\; O(G^2) \quad . \tag{103}$$

Inserting (99) and (101) into the RG equations (96) and (97), calculating all frequency integrations and neglecting all terms of $O(G^4)$, one obtains a closed set of RG equations for $G_{12}(E)|_{\bar{\omega}_1 = \bar{\omega}_2 = 0}$ and $L(E)$, which can be easily solved numerically. These constitute the basic equations of the E-RTRG approach. The crucial step in the formalism is the parametrization of the frequency dependence, otherwise a numerical solution would be very time consuming. Truncating the RG equations at $O(G^2)$ provides the solution up to leading order, whereas a truncation at $O(G^3)$ includes in addition all sub-leading terms. An important check for the reliability of the solution is whether these two truncation schemes lead approximately to the same universal solution. For the nonequilibrium Kondo model at zero magnetic field, the equations have been solved in Ref. [6] to calculate the stationary conductance with reliable results even in the strong coupling regime. Similiar RG equations can be set up for the IRLM and the spin boson model which have been studied in Refs. [7, 10].

RG equations for the slowly varying parts of the Liouvillian. We are now ready to show how the decomposition (4) can be derived together with RG equations for the slowly varying functions $L_\Delta(E)$ and $L'(E)$. First of all, one can see from the RG equations (96) and (97) that $\frac{\partial}{\partial E} L(E)$ and $G_{12}(E)$ are slowly varying logarithmic functions. At large E we find from dimensional arguments that

$$\frac{\partial^2}{\partial E^2} L(E)\,,\; \frac{\partial}{\partial E} G_{12}(E) \;\sim\; \frac{1}{E} \left(1 \,+\, O(\frac{\Delta}{E}) \right) \quad , \tag{104}$$

where Δ is some physical scale except E. For large E, we can neglect the higher orders $\sim O(\frac{\Delta}{E})$ and we see that, due to the factor $\frac{1}{E}$, logarithmic functions are generated by integrating over E. For E close to some branching point z_n, we find, that even in the worst case when all resolvents contain the same branching point, that $\frac{\partial^2}{\partial E^2} L(E)$ and $\frac{\partial}{\partial E} G_{12}(E)$ can at most diverge $\sim \frac{1}{E - z_n}$ for $E \to z_n$. As a result, also for $E \to z_n$, $\frac{\partial}{\partial E} L(E)$ and $G_{12}(E)$ are slowly varying logarithmic functions of $E - z_n$. This can only be the case if $L(E)$ consists of terms

$$L(E) \;\sim\; (E - z_n)\, K_n(E - z_n) \;=\; -z_n\, K_n(E - z_n) \,+\, E\, K_n(E - z_n) \quad , \tag{105}$$

where $K(E)$ is a slowly varying function, or, more precisely, $z_n = z_k^p - \bar{\mu}_M$ will be replaced by $\lambda_k(E_M) - \bar{\mu}_M$ if E is not close to one of the singularities. Therefore, we see that $L(E)$ can be decomposed in the form (4),

$$L(E) \;=\; L_\Delta(E) \,+\, E\, L'(E) \quad , \tag{106}$$

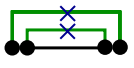
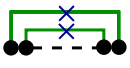
with slowly varying functions $L_\Delta(E) \sim -z_n K_n(E-z_n)$ and $L'(E) \sim K_n(E-z_n)$. We note that we used precisely this form at the end of Section 4 in Eq. (94). It shows that $L_\Delta(E)$ and $L'(E)$ have a quite similiar structure.

We note that the property that $\frac{\partial}{\partial E}L(E)$ and $G_{12}(E)$ are slowly varying logarithmic functions can also be seen directly from the original perturbative expansion (83) since in all orders of perturbation theory the number of frequency integrations is identical to the number of resolvents. This leads to logarithmic integrals at large and low energies even if all resolvents contain the same cutoff scale at low energies. For the proof it is essential that the perturbation theory is taken in the self-consistent form (83) since this leads to the property that all resolvents involve a different combination of the frequencies. The same can be shown for the ohmic spin boson model where the decomposition (4) holds also in all orders of perturbation theory. For models with 1-point vertices and a flat spectral function (like e.g. quantum dot models in the charge fluctuation regime), the number of resolvents can be arbitrarily larger than the number of frequency integrations. Here, to show the logarithmic scaling at low energies in all orders of perturbation theory, it is very important that the resolvents do not only have different frequency combinations but many of them have also different cutoff scales at low energies. In contrast to models with spin/orbital fluctuations, it turns out that already the first derivative $\frac{\partial}{\partial E}L(E)$ exists in the limit $D \to \infty$, see e.g. the first diagram of (95). This means that $\frac{\partial}{\partial E}L(E) \sim \frac{\Gamma}{E-z_n}$ multiplied with a well-controlled series with no divergence at high or low energies. This part influences only the function $L_\Delta(E)$ but not $EL'(E)$. The systematic treatment of all orders in the tunneling for models with charge fluctuations is still an issue of ongoing research.

To find RG equations for $L_\Delta(E)$ and $L'(E)$, we try to bring the RG equation (96) for $\frac{\partial^2}{\partial E^2}L(E)$ into the form

$$\frac{\partial^2}{\partial E^2}L(E) \;=\; \frac{\partial}{\partial E}L'(E) \;+\; \frac{\partial}{\partial E}\left\{\frac{\partial}{\partial E}L_\Delta(E) \;+\; E\frac{\partial}{\partial E}L'(E)\right\} \quad , \tag{107}$$

such that $\frac{\partial}{\partial E}L_\Delta(E)$ and $\frac{\partial}{\partial E}L'(E)$ can be identified and that $L_\Delta(E)$ is proportional to some physical scale Δ except E. For simplicity we show the procedure only up to $O(G^2)$, for $O(G^3)$ see Ref. [26]. Taking only the first term on the r.h.s. of the RG equation (96), replacing the vertices by the ones at zero frequency via (99), and shifting the two derivatives of the resolvent $\frac{\partial^2}{\partial E^2}R(E_{12}+\bar\omega_{12}) = \frac{\partial}{\partial\bar\omega_1}\frac{\partial}{\partial\bar\omega_2}R(E_{12}+\bar\omega_{12})$ via two partial integrations to the contractions, we obtain

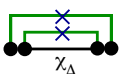
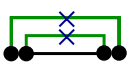
$$\frac{\partial^2}{\partial E^2}L(E) \quad = \quad \frac{1}{2}$$  $$+ \quad \frac{\partial}{\partial E}\left\{\frac{1}{2}\right.$$  $$\left.\right\} \quad + \quad O(G^3) \quad , \tag{108}$$

where a cross at a contraction denotes the derivative $\frac{\partial}{\partial\bar\omega}\gamma_1^a = \frac{\partial}{\partial\bar\omega}f_\alpha(\bar\omega)$, see (98). The dashed line in the second term indicates that the resolvent is replaced by the Z'-factor $R(E_M+\bar\omega_M) \to Z'(E_M+\bar\omega)$, defined in (29). Therefore, this term is of $O(G^3)$ and can be added without violating the consistency of the truncation scheme up to $O(G^2)$. The term has been added in such a way that when identifying (108) with (107), the derivative $\frac{\partial}{\partial E}L_\Delta(E)$ will become proportional to a physical scale Δ. Together with the relation

$$Z'(E_M+\bar\omega_M) \;-\; E\,R(E_M+\bar\omega_M) \;=\; \chi_\Delta(E,\bar\mu_M+\bar\omega_M)\,R(E_M+\bar\omega_M) \quad , \tag{109}$$

$$\chi_\Delta(E,\bar\mu_M+\bar\omega_M) \;=\; \bar\mu_M \;+\; \bar\omega_M \;-\; \tilde L_\Delta(E_M+\bar\omega_M) \quad , \tag{110}$$

which follows from (29) with the definition $\tilde{L}_\Delta(E) = Z'(E)L_\Delta(E)$, we obtain

$$\frac{\partial}{\partial E}L_\Delta(E) \;=\; \frac{1}{2}$$  $$+\; O(G^3) \quad , \quad \frac{\partial}{\partial E}L'(E) \;=\; \frac{1}{2}$$  $$+\; O(G^3)\,, \quad (111)$$

where the symbol χ_Δ at the resolvent means that the resolvent multiplied with $\chi_\Delta(E, \bar{\mu}_M + \bar{\omega}_M)$ has to be taken. Obviously, χ_Δ is proportional to a physical scale, since $\bar{\mu}_M, \bar{\omega}_M$ and $\tilde{L}_\Delta(E_M + \bar{\omega})$ have this property. For $\bar{\omega}_M$ this follows from the fact that the RG equations contain only the derivatives $\frac{\partial}{\partial\bar{\omega}}\gamma_1^a = \frac{\partial}{\partial\bar{\omega}}f_\alpha(\bar{\omega})$ of the contractions, such that $|\bar{\omega}_i| \lesssim T_{\alpha_i}$. In contrast to the RG equation (96) for the full Liouvillian $L(E)$, the RG equations (111) for $L_\Delta(E)$ and $L'(E)$ are first order differential equations. Therefore, the differences $L_\Delta(E_M + \bar{\omega}_M) - L_\Delta(E_M)$ and $L'(E_M + \bar{\omega}_M) - L'(E_M)$ are of $O(G^2)$ such that the frequency dependence of the resolvent and χ_Δ entering the RG equations (111) can be approximated by

$$R(E_M + \bar{\omega}_M) \;=\; \frac{1}{\bar{\omega}_M \;+\; E_M \;-\; \tilde{L}_\Delta(E_M)}\, Z'(E_M) \;+\; O(G^2) \quad (112)$$

$$\chi_\Delta(E, \bar{\mu}_M + \bar{\omega}_M) \;=\; \bar{\mu}_M \;+\; \bar{\omega}_M \;-\; \tilde{L}_\Delta(E_M) \;+\; O(G^2) \quad . \quad (113)$$

As a consequence, all frequency integrations can be straightforwardly performed such that the differential-integro equations (111) are converted into differential equations. E.g., at zero temperature, the two frequency integrations in (111) are trivial leading to the explicit expression

$$\frac{\partial}{\partial E}L_\Delta(E) \;=\; \frac{1}{2}\,G_{12}(E)\,\frac{\bar{\mu}_{12} \;-\; \tilde{L}_\Delta(E_{12})}{E_{12} \;-\; \tilde{L}_\Delta(E_{12})}\,Z'(E_{12})\,G_{\bar{2}\bar{1}}(E_{12}) \quad , \quad (114)$$

$$\frac{\partial}{\partial E}L'(E) \;=\; \frac{1}{2}\,G_{12}(E)\,\frac{1}{E_{12} \;-\; \tilde{L}_\Delta(E_{12})}\,Z'(E_{12})\,G_{\bar{2}\bar{1}}(E_{12}) \quad , \quad (115)$$

together with the RG equation for the vertex which follows from the lowest order term of (97) as

$$\frac{\partial}{\partial E}G_{12}(E) \;=\; G_{13}(E)\,\frac{1}{E_{13} \;-\; \tilde{L}_\Delta(E_{13})}\,Z'(E_{13})\,G_{\bar{3}2}(E_{13}) \;-\; (1 \leftrightarrow 2) \quad . \quad (116)$$

Solution of approximate E-RTRG equations. The first-order RG equations (114) and (115) for $L_\Delta(E)$ and $L'(E)$ provide the most convenient starting point for an analytical solution of the RG equations at least in that regime of the complex plane where the effective vertices stay small, see Refs. [10, 26] for details. The strategy is to solve the RG equations approximately in three different energy regimes by expanding in the effective vertices but keeping large logarithmic terms (either at large or low energies) to all orders, and matching the different solutions to fix the integration constants. Denoting the small dimensionless coupling constant by α, we distinguish the following regimes: (1) The regime of high energies $|E| \gg |z_n|$, where the RG resums all ultraviolett logarithmic terms $\sim (\alpha \ln \frac{D}{-iE})^k$; (2) The regime of intermediate and small energies $|E - z_n| \lesssim O(|z_n|)$ but E not too close to the branching points such that one can expand in the small parameter $\alpha|\ln \frac{|z_n|}{|E-z_n|}| \ll 1$; (3) The regime of small energies exponentially close to some of the branching points, i.e. $|E - z_n| \ll O(|z_n|)$ and $\alpha|\ln \frac{|z_n|}{|E-z_n|}| \sim 1$, where the RG resums all infrared logarithmic terms $\sim (\alpha|\ln \frac{|z_n|}{|E-z_n|}|)^k$. In particular for the ohmic spin boson model and the IRLM, we will see in Section 6 that the coupling constant α stays small in the whole complex plane such that a well-controlled analytical solution is possible for

all E, showing that the resummation of logarithmic terms for high and low energies gives very different results. For the Kondo problem a weak-coupling solution is only possible for high, intermediate and small energies, but not for exponentially small energies where the coupling constant $\alpha \sim O(1)$.

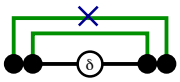
Initial conditions. The initial conditions for the RG flow at large energies are set up at the value $E = iD$, where $D \gg |z_n|$ is the high-energy cutoff. The motivation for the choice $E = iD$ lies in the fact that, for $D \gg |E| \gg |z_n|$, the bare perturbation series for $L_\Delta(E)$ and $L'(E)$ contain logarithmic terms $\sim (\ln \frac{D}{-iE})^k$ of all powers k (we have chosen $-iE$ in the argument, such that the branch cut is directed towards the negative imaginary axis). All other terms $\sim (\frac{|E|}{D})^n$ are neglected since they vanish in the limit $D \to \infty$ and thus do not contribute to the universal solution which is independent of the cutoff details. Extrapolating this result up to $E = iD$ has the effect that all logarithmic terms vanish, which sets the initial point for the universal RG flow. The calculation of the initial values can be done by bare perturbation theory for $D \gg |E| \gg |z_n|$ and omitting all logarithmic contributions. For small bare coupling constants it is sufficient to take the lowest order term if it is universal, otherwise one takes zero for the initial condition. We note that this procedure works well to determine the universal initial condition for $L_\Delta(E)$ and $L'(E)$ at $E = iD$ but fails for the initial condition of $L(E)$ since $L(E)$ contains terms linear in E which are very large for $E = iD$. Therefore, for the RG equation (96) one either has to keep the high-energy cutoff function in the RG equations and start the RG flow at $|E| \gg D$ (where one can take the bare values as initial condition), or one has to find a reference point at low energies where $L(E)$ is known from exact results, see e.g. the solution of the Kondo model in strong coupling in Ref. [6].

RG for the Liouvillian discontinuity jumps. Finally we show how RG equations can be derived for the jump δL of the Liouvillian at a particular branch cut, as this jump is needed to evaluate the branch cut contributions to the time evolution, see Eq. (48). As described in Section 4 the branch cuts of $L(E)$ occur only at zero temperature and can be identified in the perturbative expansion (83) by closing all frequency integrations in the upper half of the complex plane and considering the branch cuts of the Fermi distribution functions on the positive imaginary axis. This means that the frequencies are shifted to the positive imaginary axis $\bar\omega_M \to i|\bar\omega_M|$. In leading order, a given branch cut at $E = z_n - ix \pm 0^+$ is generated by some resolvent which is resonant, i.e. the jump of this resolvent across the branch cut becomes a δ-function. If the resolvent contains the eigenvalue $\lambda_k(E_M + i|\bar\omega_M|)$ the resonance occurs if $z_n = z_k^p - \bar\mu_M$. With $E_M = z_n + \bar\mu_M - ix \pm 0^+ = z_k^p - ix \pm 0^+$, we replace approximately $\lambda_k \to z_k^p$, $P_k \to \bar P_k(z_k^p - ix)$ and $Z' \to \bar Z'(z_k^p - ix)$, which gives for the jump of the resolvent the following δ-function

$$\begin{aligned} \delta R &= \left(\frac{1}{-ix + i\bar\omega_{1\ldots n} + 0^+} - \frac{1}{-ix + i|\bar\omega_M| - 0^+} \right) \bar P_k(z_k^p - ix)\, \bar Z'(z_k^p - ix) \\ &= 2\pi\delta(|\bar\omega_M| - x)\, \bar P_k(z_k^p - ix)\, \bar Z'(z_k^p - ix) \quad . \end{aligned} \tag{117}$$

As expected, the frequency integrals give only a contribution for $x > 0$, since this is the region where the branch cut starts. The RG equation for $\frac{\partial \delta L}{\partial E}(z_n - ix)$ is obtained by fixing the resonant resolvent together with the resolvent where the E-derivative is taken and resumming the rest of the perturbative series in terms of effective 2-point vertices. If both resolvent are the same, the E-derivative is replaced by a frequency derivative $\frac{\partial}{\partial \bar\omega_i}$ and is shifted via partial integration to

the derivative of some contraction crossing over the resolvent. Thus, we obtain in leading order

$$\frac{\partial \delta L}{\partial E}(z_n - ix) \;=\; i\frac{\partial}{\partial x}\delta L(z_n - ix) \;=\; -\frac{1}{2}$$  (118)

where the symbol δ at the resolvent means that we replace the resolvent by its jump δR given by (117). The initial condition for the RG equation is $\delta L(z_n) = 0$. At zero temperature both frequency integrations are trivial. The contraction with the cross gives $\int d\bar{\omega} f'_\alpha(\bar{\omega})\{\dots\} = -\{\dots\}_{\bar{\omega}=0}$, whereas the contraction without the cross leads to

$$\int d\bar{\omega}\,(f_\alpha(\bar{\omega}) - \frac{1}{2})\,\{\dots\} \;=\; -\frac{1}{2}\int d\bar{\omega}\,\mathrm{sign}(\bar{\omega})\,\{\dots\} \;=\; -i\int_0^\infty d\bar{\omega}\,\{\dots\}_{\bar{\omega}\to i\bar{\omega}} \quad , \quad (119)$$

where we have used that the jump of the sign-function at the branch cut is given by 2. Therefore, (118) gives explicitly

$$\frac{\partial}{\partial x}\delta L(z_n - ix) \;=\; -\pi\,\theta(x)\,\bar{G}_{12}(z_k^p - \bar{\mu}_{12} - ix)\,\bar{P}_k(z_k^p - ix)\,\bar{Z}'(z_k^p - ix)\,\bar{G}_{\bar{2}\bar{1}}(z_k^p - ix) \quad , \quad (120)$$

where we have replaced all vertices that are discontinous across the branch cut by their average value $\bar{G}$. Up to the corrections from the weak x-dependence of the logarithmic functions on the r.h.s. of this equation, we obtain

$$\delta L(z_n - ix) \;\sim\; x\,\theta(x) \quad , \quad (121)$$

giving Eq. (50) used in Section 3. Therefore, if x can be neglected in the resolvents of the integrand of (88), we obtain (up to logarithmic corrections) $\rho_t^{n,b} \sim 1/t^2$ in the long-time limit for *all* models with spin or orbital fluctuations, like e.g. the Kondo model. Similiar considerations show that the same holds for the ohmic spin boson model, whereas for models with charge fluctuations (like the IRLM), one obtains $\delta L(z_n - ix) \sim \theta(x)$.

6 Results

In this section we will discuss the application of the formalism to the models introduced in Section 2. Since the Kondo model has the simplest algebra, we will take this model as a tutorial example to discuss the solution of the RG equations and the consequences for the time evolution in all detail. Since the general strategy is always the same, we will then show briefly the results for the ohmic spin boson model and the IRLM, and will concentrate on interesting features which are different from the ones for the Kondo model.

6.1 Kondo model

We consider the nonequilibrium Kondo model at zero magnetic field $h^{(0)} = 0$ and zero temperature $T = 0$ for the antiferromagnetic case $J^{(0)} > 0$. We assume that the local spin-$\frac{1}{2}$ is coupled to several reservoirs with chemical potentials μ_α. For the special case of two reservoirs $\alpha \equiv L/R \equiv \pm$, we take $\mu_\alpha = \alpha\frac{V}{2}$, where V denotes the bias voltage across the system. Following Refs. [5, 6] our aim is to calculate the time evolution of the local spin $\langle \underline{S} \rangle(t)$.

The model is spin rotational invariant and therefore the effective Liouvillian $L(E)$ should be an invariant under spin rotations. Defining two basis spinoperators $\underline{L}_{\pm}$ in Liouville space by (A is an arbitrary local operator) $\underline{L}^+ A = \underline{S} A$ and $\underline{L}^- A = -A\underline{S}$, the only two invariants are given by the identity and $\underline{L}^+ \cdot \underline{L}^-$. Since the Liouvillian must also fulfil $\mathrm{Tr} L(E) = 0$, we find that the Liouvillian can be parametrized by

$$L(E) \;=\; -i\,\Gamma(E)\,L^a \quad , \quad L^a \;=\; \frac{3}{4} \;+\; \underline{L}^+ \cdot \underline{L}^- \quad . \tag{122}$$

$\Gamma(E)$ is the energy dependent spin relaxation rate. The Liouvillian has one zero eigenvalue with projector $P_{\rm st} = 1 - L^a$ and three degenerate eigenvalues at $-i\Gamma(E)$ with projector $1 - P_{\rm st} = L^a$. Therefore, by using (28), we can write for the time evolution of the local density matrix

$$\rho(t) \;=\; \rho_{\rm st} \;+\; \frac{i}{2\pi} \int_{-\infty+i0^+}^{\infty+i0^+} dE\, e^{-iEt}\, \frac{1}{E \;+\; i\Gamma(E)}\, L^a\, \rho_{t=0} \quad , \tag{123}$$

where $\rho_{\rm st} = (1 - L^a)\rho_{t=0}$ is the diagonal stationary density matrix with equal probabilities for both spin directions $(\rho_{\rm st})_{ss} = \frac{1}{2}$. Using $\mathrm{Tr}\underline{S}\rho(t) = \mathrm{Tr}\underline{S}L^a\rho(t)$ and $(L^a)^2 = L^a$, we find that the spin relaxation rate $\Gamma(E)$ determines the spin dynamics via

$$\langle \underline{S} \rangle(t) \;=\; \frac{i}{2\pi} \int_{-\infty+i0^+}^{\infty+i0^+} dE\, e^{-iEt}\, \frac{1}{E \;+\; i\Gamma(E)}\, \langle \underline{S} \rangle_{t=0} \quad . \tag{124}$$

As a consequence, the operator structure of the Liouvillian is no longer important, and we can use all formulas derived in Section 28 for the time evolution with the replacement $L(E) \to -i\Gamma(E)$ or $R(E) \to \frac{1}{E+i\Gamma(E)}$. This means that all projectors P_k can be left out and $Z'(E)$ can be used from the decomposition

$$\Gamma(E) \;=\; \Gamma_\Delta(E) \;+\; E\,\Gamma'(E) \quad , \quad R(E) \;=\; \frac{1}{E \;+\; i\Gamma(E)} \;=\; \frac{1}{E \;+\; i\tilde{\Gamma}_\Delta(E)}\, Z'(E) \quad , \tag{125}$$

with

$$\tilde{\Gamma}_\Delta(E) \;=\; Z'(E)\,\Gamma_\Delta(E) \quad , \quad Z'(E) \;=\; \frac{1}{1 \;+\; i\Gamma'(E)} \quad . \tag{126}$$

Analysing the functions $\tilde{\Gamma}_\Delta(E)$ and $Z'(E)$ in leading order from the RG equations (114) and (115), we will show below that the branching poles and branching points of the resolvent $R(E)$ are given by

$$z_0^p \;=\; -i\Gamma^* \quad , \quad z^b_{\alpha\alpha'} \;=\; -i\Gamma^* \;+\; \mu_\alpha \;-\; \mu_{\alpha'} \quad \text{for} \quad \alpha \neq \alpha' \tag{127}$$

where we used the notation of Eq. (41) and assumed that all reservoirs have different chemical potentials (otherwise they can be taken together). $z_0^p = i\tilde{\Gamma}_\Delta(z_0^p)$ is the pole of the resolvent $R(E)$ and Γ^* is called the Korringa rate which, as will be shown below, is given by

$$\Gamma^* \;=\; 2\pi\, J_V^2 \sum_{\alpha\neq\alpha'} x_\alpha\, x_{\alpha'}\, |\mu_\alpha - \mu_{\alpha'}| \quad , \quad J_V \;=\; \frac{1}{2\,\ln\frac{V}{T_K}} \quad , \quad T_K \;=\; D\, e^{-1/(2J^{(0)})} \quad , \tag{128}$$

where we have used the notation of Eq. (13). J_V is the renormalized exchange coupling at the scale $V = \max_{\alpha\alpha'}\{|\mu_\alpha - \mu_{\alpha'}|\}$ (which is the bias voltage for two reservoirs) and T_K is called the Kondo temperature . The result for the Korringa rate holds in the weak-coupling case $V \gg T_K$

which we will consider from now on. In this case, we get $J_V \ll 1$ and $\Gamma^* \ll V$. Using (41) and (127), we get the following general form for the time evolution of the local spin

$$\langle \underline{S} \rangle(t) \;=\; F_0^p(t)\, e^{-\Gamma^* t}\, \langle \underline{S} \rangle(t)_{t=0} \;+\; \sum_{\alpha \neq \alpha'} F^b_{\alpha\alpha'}(t)\, e^{-\Gamma^* t}\, e^{-i(\mu_\alpha - \mu_{\alpha'})t}\, \langle \underline{S} \rangle(t)_{t=0} \quad . \tag{129}$$

For **intermediate and long times** $t \gtrsim \frac{1}{V}$, the pre-exponential functions can be calculated from (45) and (48) as

$$F_0^p(t) \;=\; \bar{Z}'(-i\Gamma^* - i/t) \quad , \tag{130}$$

$$F^b_{\alpha\alpha'}(t) \;=\; -\frac{i}{2\pi} \int_0^\infty dx\, e^{-xt}\, \frac{\bar{Z}'(z^b_{\alpha\alpha'} - i/t)^2\, \delta\Gamma(z^b_{\alpha\alpha'} - ix)}{(\mu_\alpha \,-\, \mu_{\alpha'} \,-\, ix)^2} \quad , \tag{131}$$

where we have neglected $-i\Gamma^* + i\bar{\tilde{\Gamma}}_\Delta(z^b_{\alpha\alpha'} - ix)$ in the dominator of the integrand of the last equation. This can be done for $|\mu_\alpha - \mu_{\alpha'}| \gg \Gamma^*$ for all $\alpha \neq \alpha'$, i.e. if the branch cuts are sufficiently apart from each other (other cases can be treated as well but need a special procedure [5]). For **short times** $t \ll \frac{1}{V}$, we can use (42) and get

$$\langle \underline{S} \rangle(t) \;=\; Z'(1/t)\, e^{-\tilde{\Gamma}_\Delta(1/t)t}\, \langle \underline{S} \rangle(t)_{t=0} \quad . \tag{132}$$

To evaluate (130), (131) and (132) explicitly, we need the functions $\tilde{\Gamma}_\Delta(E)$, $Z'(E)$ and $\delta\Gamma(z_{\alpha\alpha'} - ix)$, which we will derive in the following by considering the RG equations (114), (115) and (120), together with the RG equation (116) for the vertex.

In leading order it can be shown that the effective vertex $G_{11'}(E)$ can be parametrized in the same form as the initial vertex $G^{(0)}_{11'}$, defined by (15) and (55). This gives the form

$$G_{11'}(E) \;=\; -\, J_{\alpha\alpha'}(E)\, \underline{L}^2 \cdot \underline{\sigma}_{\sigma\sigma'} \quad \text{for} \quad \eta = -\eta' = + \quad , \quad \underline{L}^2 \;=\; -\frac{1}{2}\,(\underline{L}^+ \,+\, \underline{L}^-) \quad , \tag{133}$$

together with $G_{11'}(E) = -G_{1'1}(E)$ for $\eta = -\eta' = -$. Using this ansatz together with the form (122) for the Liouvillian in the RG equations and omitting the terms $\sim L^a$ in the RG equation for the vertex (which generate higher orders), one finds after some straightforward algebra [6] the following RG equations

$$\frac{\partial}{\partial E}\Gamma_\Delta(E) \;=\; i\, \chi_\Delta(E, \hat{\mu}_{\alpha\alpha'})\, R(\hat{E}_{\alpha\alpha'})\, J_{\alpha\alpha'}(E)\, J_{\alpha'\alpha}(\hat{E}_{\alpha\alpha'}) \quad , \tag{134}$$

$$\frac{\partial}{\partial E} Z'(E) \;=\; Z'(E)^2\, R(\hat{E}_{\alpha\alpha'})\, J_{\alpha\alpha'}(E)\, J_{\alpha'\alpha}(\hat{E}_{\alpha\alpha'}) \quad , \tag{135}$$

$$\frac{\partial}{\partial E} J_{\alpha\alpha'}(E) \;=\; -\frac{1}{2}\, R(\hat{E}_{\alpha\alpha''})\, J_{\alpha\alpha''}(E)\, J_{\alpha''\alpha'}(\hat{E}_{\alpha\alpha''}) \;-\; \frac{1}{2}\, R(\hat{E}_{\alpha''\alpha'})\, J_{\alpha''\alpha'}(E)\, J_{\alpha\alpha''}(\hat{E}_{\alpha''\alpha}) \quad , \tag{136}$$

where $\hat{E}_{\alpha\alpha'} = E + \hat{\mu}_{\alpha\alpha'}$, $\hat{\mu}_{\alpha\alpha'} = \mu_\alpha - \mu_{\alpha'}$, $\chi_\Delta(E, \hat{\mu}_{\alpha\alpha'}) = \hat{\mu}_{\alpha\alpha'} + i\tilde{\Gamma}_\Delta(\hat{E}_{\alpha\alpha'})$, and $Z'(E)$ and $R(E)$ have been defined in (125) and (126). The initial conditions at $E = iD$ are $\Gamma_\Delta = 0$, $Z' = 1$ and $J_{\alpha\alpha'} = 2\sqrt{x_\alpha x_{\alpha'}} J^{(0)}$.

We first start with the analytic solution in the regime of **high energies** $|E| \gg V$. Neglecting $\hat{\mu}_{\alpha\alpha'}$ everywhere gives the solution $\Gamma_\Delta(E) \approx 0$ and $J_{\alpha\alpha'}(E) \approx 2\sqrt{x_\alpha x_{\alpha'}} J(E)$, together with

$$\frac{\partial}{\partial E} Z'(E) \;=\; \frac{4}{E}\, Z'(E)^2\, J(E)^2 \;=\; \frac{4}{E}\, J(E)^2 + O(J^3) \quad , \tag{137}$$

$$\frac{\partial}{\partial E} J(E) \;=\; -\frac{2}{E}\, Z'(E)\, J(E)^2 \;=\; -\frac{2}{E}\, J(E)^2 + O(J^3) \quad , \tag{138}$$

where we have used $Z' = 1 + O(J)$ on the r.h.s. (see Eq. (139)). We find that $-\frac{1}{2J(E)} + \ln(-iE) \equiv \ln T_K$ is an invariant and $\frac{\partial Z'}{\partial J} = -2$, which gives the solution

$$Z'(E) \;=\; 1 \,-\, 2\,(J(E) \,-\, J^{(0)}) \,\to 1 \,-\, 2\,J(E) \quad , \quad J(E) \;=\; \frac{1}{2\ln\frac{-iE}{T_K}} \quad , \tag{139}$$

where the Kondo temperature T_K is defined in (128) and we used the scaling limit in the first equation, defined by $D \to \infty$, $J^{(0)} \to 0$, such that T_K remains a constant. We have chosen $-iE$ in the argument of the logarithm to define a real value for the Kondo temperature at $E = iD$ and since we want the branch cut of the logarithm to point into the direction of the negative imaginary axis. In the solution (139) all logarithmic terms $\sim (J^{(0)} \ln\frac{D}{-iE})^k$ have been resummed. $J(E)$ is the poor man scaling solution of the Kondo model, already introduced in the lecture B3 by T. Costi, but with the difference that $-iE$ plays now the role of the effective energy scale. Most importantly, the solution would diverge at $E = iT_K$ when extrapolated to small energies, indicating an increase of antiferromagnetic spin fluctuations at the scale of the Kondo temperature. However, in this regime the solution can not be used since the RG flow becomes very different for $|E| \lesssim V$. The solution at high energies can be used to evaluate the universal solution (132) for **short times** $t \ll \frac{1}{V}$

$$\langle \underline{S} \rangle(t) \;=\; \left(1 \,-\, \frac{1}{|\ln(T_K t)|}\right) \langle \underline{S} \rangle(t)_{t=0} \quad . \tag{140}$$

In this result all logarithmic terms $\sim (J^{(0)} \ln(Dt))^k$ have been resummed, which can be seen from $1/\ln(T_K) = -2J^{(0)}/(1 - 2J^{(0)} \ln(Dt))$. Sub-leading terms are not included but can be taken into account by truncating the RG equations at $O(G^3)$ [5]. For ferromagnetic Kondo models the universal short time behaviour has also been discussed in Ref. [27] using flow equation methods.

To find the solution at **intermediate and small energies** $|E| \lesssim V$ but not exponentially close to the singularites such that $J(E) \ll 1$ is still fulfilled (we state below what this precisely means), we first set the initial value by expanding the solution (139) at high energies for the case when E starts to approach V such that $|E| \gg V$ is still fulfilled but $J_V |\ln\frac{-iE}{V}| \ll 1$, where $J_V = 1/(2\ln(V/T_K))$ is the exchange coupling at high energies evaluated at the scale $-iE = V$, as introduced in (128). This gives

$$J_{\alpha\alpha'}(E) \;\approx\; 2\sqrt{x_\alpha x_{\alpha'}}\, J_V\, (1 \,-\, 2\, J_V\, \ln\frac{-iE}{V}) \quad , \quad Z'(E) \;\approx\; 1 \,-\, 2\, J_V \,+\, 4\, J_V^2\, \ln\frac{-iE}{V} \quad . \tag{141}$$

Using the first term of this expansion in the r.h.s. of the full RG equations (134), (135) and (136),

we can easily integrate the equations up to $O(J_V^2)$ by using $R(E) = \frac{\partial}{\partial E} \ln \frac{-i(E+i\tilde{\Gamma}_\Delta(E))}{V} + O(J_V^2)$

$$\tilde{\Gamma}_\Delta(E) \;=\; i\,4\,x_\alpha x_{\alpha'}\,J_V^2\,(\hat{\mu}_{\alpha\alpha'} + i\tilde{\Gamma}_\Delta(\hat{E}_{\alpha\alpha'}))\,\ln\frac{-i(\hat{E}_{\alpha\alpha'} + i\tilde{\Gamma}_\Delta(\hat{E}_{\alpha\alpha'}))}{V} \quad , \tag{142}$$

$$Z'(E) \;=\; 1 \;-\; 2\,J_V \;+\; 4\,x_\alpha x_{\alpha'}\,J_V^2\,\ln\frac{-i(\hat{E}_{\alpha\alpha'} + i\tilde{\Gamma}_\Delta(\hat{E}_{\alpha\alpha'}))}{V} \quad , \tag{143}$$

$$J_{\alpha\alpha'}(E) \;=\; 2\sqrt{x_\alpha x_{\alpha'}}\,J_V\,\Big\{1 \;-\; x_{\alpha''}\,J_V\,\ln\frac{-i(\hat{E}_{\alpha\alpha''} + i\tilde{\Gamma}_\Delta(\hat{E}_{\alpha\alpha''}))}{V} - x_{\alpha''}\,J_V\,\ln\frac{-i(\hat{E}_{\alpha''\alpha'} + i\tilde{\Gamma}_\Delta(\hat{E}_{\alpha''\alpha'}))}{V}\Big\} \quad . \tag{144}$$

The integration constants have been chosen such that for $|E| \gg V$ and $J_V|\ln\frac{-iE}{V}| \ll 1$, the result (141) at high energies is reproduced. From the solution we can see that branch cuts appear starting at the singularities $z_n = -i\Gamma^* + \hat{\mu}_{\alpha\alpha'}$, as stated in (127). Furthermore, we can see that the expansion is well-defined provided that $J_V|\ln\frac{V}{|E-z_n|}| \ll 1$, which is the precise condition that E should not be exponentially close to the branching points. This is the reason why the scale $V \sim |z_n|$ has been chosen as reference scale in the logarithm to integrate the RG equations perturbatively for intermediate and small energies. In the solution all logarithmic terms $\sim (J^{(0)}\ln\frac{D}{V})^k$ have been resummed in J_V, whereas a perturbative treatment has been used for the logarithmic terms $J_V|\ln\frac{V}{|E-z_n|}| \ll 1$.

Since $\tilde{\Gamma}_\Delta(E)$ is a weakly varying function for $|E| \lesssim |z_n|$, we can replace $\tilde{\Gamma}_\Delta(\hat{E}_{\alpha\alpha'}) \to \Gamma^*$ in the above equations and neglect the term $\sim J_V^2\Gamma^* \sim J_V^4$ in (142). Inserting $E = -i\Gamma^* + \delta$ in (142) (where $|\delta| \ll \Gamma^*$ is a small scale to exclude an exponentially small region around $z_0^p = -i\Gamma^*$), we find straightforwardly the result (128) for Γ^*. Inserting the solution for $Z'(E)$ in (130), we can calculate the pre-exponential function for the contribution from the branching pole at $E = z_0^p = -i\Gamma^*$. For **long times** $t \gg \frac{1}{V}$ we obtain

$$F_0^p(t) \;=\; 1 \;-\; 2\,J_V \;-\; 4\sum_\alpha x_\alpha^2\,J_V^2\,\ln(Vt) \quad , \tag{145}$$

whereas, for intermediate times $t \sim \frac{1}{V}$, the contribution in $O(J_V^2)$ is not logarithmic and unimportant (the precise coefficient is also influenced by other sub-leading terms). Several interesting features appear in this result. The first term is the result from a Markov approximation, where only the pole without the residuum is considered. We note that the pole position is also influenced by non-Markovian contributions arising when (142) is solved self-consistently for Γ^*. Here, this is a very weak effect occuring in $O(J_V^4)$. For quantum dot models such non-Markovian contributions have been dicussed perturbatively in Ref. [28]. All other terms of (145) are of pure non-Markovian nature arising from the term linear in E of the effective Liouvillian (leading to the Z'-factor). The second term linear in J_V can not be obtained from perturbation theory since this term of the Z'-factor involves the difference $J_V - J^{(0)}$ (see Eq. (139)), which reduces to J_V only in the scaling limit. It arises from a resummation of a series of logarithmic terms $\sim (J^{(0)}\ln\frac{D}{V})^k$ which starts at $k = 2$, i.e. the $k = 1$ term is absent. The last term $\sim J_V^2\ln(Vt)$ is logarithmic in time and becomes of $O(1)$ for exponentially large times $t \sim \frac{1}{V}e^{1/J_V^2}$. In this regime the solution can no longer be used since it corresponds to the regime of energies E exponentially close to z_0^p. In this regime, the full RG equation (136) for the vertex

shows that $J(E)$ does not stay small, i.e. a strong coupling problem arises and the truncation scheme is no longer controlled. As a consequence we see that, concerning the long-time evolution at exponentially large times, even in the regime $V \gg T_K$, a strong coupling method is needed to calculate pre-exponential functions. On the other hand, the exponential decay $e^{-\Gamma^* t}$ leads to a very small contribution for exponentially large times, so that it is of no practical use to know the pre-exponential function in this regime. However, for other problems with quantum critical points, like e.g. multi-channel Kondo models or the sub-ohmic spin boson model, it happens that the pole $z_0^p = 0$ lies at the origin such that no exponential decay appears. For such models, it is an interesting subject for the future to calculate the precise form of $F_0^p(t)$ for exponentially large times. E.g., for a multi-channel Kondo model with many channels $N \gg 1$, which turns out to be a weak-coupling problem in the whole complex plane, it has been shown in Ref. [6] that $F_0^p(t) \sim (\frac{1}{T_K t})^{4/N}$.

Finally, to calculate the branch cut contribution (131), we need also the jump $\delta\Gamma(z^b_{\alpha\alpha'} - ix)$ for $x \sim 1/t \lesssim V$. This is obtained from the RG equation (120), which reads with $J_{\alpha\alpha'}(E) \to 2\sqrt{x_\alpha x_{\alpha'}} J_V$

$$\frac{\partial}{\partial x}\delta\Gamma(z^b_{\alpha\alpha'} - ix) \;=\; -\,8\pi i\, x_\alpha x_{\alpha'}\, J_V^2\, \theta(x) \quad . \tag{146}$$

This leads to $\delta\Gamma(z^b_{\alpha\alpha'} - ix) = -8\pi i x_\alpha x_{\alpha'} J_V^2 x\theta(x)$. Inserting this result in (131) and using $Z' = 1 + O(J_V^2)$, we obtain for **long times** $t \gg \frac{1}{|\mu_\alpha - \mu_{\alpha'}|}$ (note that $\alpha \neq \alpha'$ and we assumed that $|\mu_\alpha - \mu_{\alpha'}| \gg \Gamma^*$)

$$F^b_{\alpha\alpha'}(t) \;=\; -\,4x_\alpha x_{\alpha'}\, J_V^2 \left(\frac{1}{(\mu_\alpha - \mu_{\alpha'})t}\right)^2 \quad . \tag{147}$$

Other time regimes $t \sim \frac{1}{|\mu_\alpha - \mu_{\alpha'}|}$ can also be studied leading to exponential integrals [5]. Inserting (147) in (129) we get an oscillating term $\sim J_V^2 (\frac{1}{(\mu_\alpha - \mu_{\alpha'})t})^2 e^{-\Gamma^* t} e^{-i(\mu_\alpha - \mu_{\alpha'})t}$ for the time evolution of the local spin. It appears in second order in J_V and is again of non-Markovian nature. In contrast to the Markov contribution it oscillates with a frequency set by the differences of chemical potentials and the pre-exponential function decays as a power law $\sim 1/t^2$ for long times. This behaviour is quite generic for models with spin or orbital fluctuations. In higher orders the oscillation frequencies are set by the renormalized excitation energies of the system associated with certain processes. E.g. a process where a particle is transferred from reservoir α to reservoir α' involves an energy cost $\mu_{\alpha'} - \mu_\alpha$, which gives the oscillation frequency. In the presence of a local magnetic field $h^{(0)}$, the same process costs the energy $\mu_{\alpha'} - \mu_\alpha \pm h$ if the local spin is flipped, where h is the renormalized magnetic field. As a consequence, these scales define further oscillation frequencies. In addition, each process has its own decay rate, setting the scale of the exponential decay. These issues have been discussed in detail in Ref. [5] for the case of the anisotropic Kondo model at finite magnetic field.

6.2 Ohmic spin boson model

Here, we consider the ohmic spin boson model at zero bias $\epsilon = 0$ and zero temperature $T = 0$. We will follow Ref. [10] where the model has recently been solved for weak damping $\alpha \ll 1$ by a systematic RG analysis using the E-RTRG method. In contrast to the Kondo model it turns out that the effective vertex $G(E) = G_1(E)_{\bar\omega=0}$ at zero frequency stays small in the whole complex plane allowing for a full solution of the problem on all time scales. We show here only the solution since the derivation is very similiar to the one for the Kondo model, except that the

algebra is more involved and the solution of the RG equations can also be derived for E close to the branching points z_n. It turns out that the resolvent $R(E)$ has four poles at

$$z_{\rm st} \,=\, 0 \quad , \quad z_0 \,=\, -i\Gamma \quad , \quad z_\pm \,=\, \pm\tilde{\Delta} \,-\, i\Gamma/2 \quad , \tag{148}$$

where

$$\tilde{\Delta} \,=\, \Delta \left(\frac{\tilde{\Delta}}{D}\right)^{\alpha} \,=\, \Delta \left(\frac{\Delta}{D}\right)^{\frac{\alpha}{1-\alpha}} \quad , \quad \Gamma \,=\, \pi\alpha\tilde{\Delta} \quad . \tag{149}$$

$\tilde{\Delta}$ is called the renormalized tunneling which is kept fixed in the scaling limit $D \,\to\, \infty$ and $\alpha \,\to\, 0$. In leading order truncation at $O(\alpha)$ it turns out that no branching poles appear, i.e. all poles are isolated. In addition, the eigenvalue $\lambda_0(E)$ has two branch cuts starting at $z_\pm$ and the eigenvalues $\lambda_\pm(E)$ have a branch cut starting at z_0. Therefore, according to the general expression (41) we get

$$\rho(t) \,=\, \rho_{\rm st} \,+\, \sum_{k=0,\pm} \left(F_k^p(t) \,+\, F_k^b(t)\right) \, e^{-iz_k t} \, \rho_{t=0} \quad , \tag{150}$$

i.e. all singularities can either act as a pole or as a branch cut. Ordering the four possible states in Liouville space by $++, --, +-, -+$, where $\pm$ are the two local states, one can show that the stationary density matrix is given by $\rho_{\rm st} \,=\, \frac{1}{2}(1, 1, \frac{\tilde{\Delta}}{\Delta}, \frac{\tilde{\Delta}}{\Delta})$ and the pre-exponential functions for **long times** $\tilde{\Delta}t \,\gg\, 1$ (note that this includes the important regime $\Gamma t \,\sim\, O(1)$ where the exponentials are of $O(1)$) are given by [10]

$$F_0^p(t) = \frac{\tilde{\Delta}}{\Delta} \begin{pmatrix} 0 & 0 \\ -1 & \tilde{\Delta}/\Delta \end{pmatrix} \otimes \tau_+ \quad , \quad F_\pm^p(t) \,=\, \frac{1}{2} \begin{pmatrix} 1 & \pm\tilde{\Delta}/\Delta \\ \pm\tilde{\Delta}/\Delta & (\tilde{\Delta}/\Delta)^2 \end{pmatrix} \otimes \tau_- \quad , \tag{151}$$

$$F_0^b(t) \,=\, -\,2\alpha \, \frac{1}{(\tilde{\Delta}t)^2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \tau_- \quad , \quad F_\pm^b(t) \,=\, -\,\alpha \, \frac{s(t)}{(\Delta t)^2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \otimes \tau_+ \quad , \tag{152}$$

where $\tau_\pm \,=\, \frac{1}{2}(1 \pm \sigma_x)$, and, for two 2×2-matrices A and B, we have defined the 4×4-matrix

$$A \otimes B \,\equiv\, \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix} \quad . \tag{153}$$

Furthermore, the logarithmic function $s(t)$ is defined by

$$s(t) \,=\, \left(\frac{1}{[1 \,+\, \alpha \ln(\tilde{\Delta}t)]\,[1 \,-\, \ln(1 \,+\, \alpha \ln(\tilde{\Delta}t))]}\right)^2 \quad . \tag{154}$$

In terms of the expectation values of the Pauli matrices $\langle\sigma_i\rangle(t) = {\rm Tr}\sigma_i\rho(t)$, these equations can also be written as

$$\begin{pmatrix} 1 \\ \langle\sigma_x\rangle(t) \end{pmatrix} = \begin{pmatrix} 1 \\ \tilde{\Delta}/\Delta \end{pmatrix} + \frac{\tilde{\Delta}}{\Delta} \begin{pmatrix} 0 & 0 \\ -1 & \tilde{\Delta}/\Delta \end{pmatrix} \begin{pmatrix} 1 \\ \langle\sigma_x\rangle_{t=0} \end{pmatrix} e^{-iz_0 t} \\ - \,\alpha \, \frac{s(t)}{(\Delta t)^2} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \langle\sigma_x\rangle_{t=0} \end{pmatrix} \sum_{\sigma=\pm} e^{-iz_\sigma t} \quad , \tag{155}$$

$$\begin{pmatrix} \langle\sigma_z\rangle(t) \\ -i\langle\sigma_y\rangle(t) \end{pmatrix} = \frac{1}{2} \sum_{\sigma=\pm} \begin{pmatrix} 1 & \sigma\tilde{\Delta}/\Delta \\ \sigma\tilde{\Delta}/\Delta & (\tilde{\Delta}/\Delta)^2 \end{pmatrix} \begin{pmatrix} \langle\sigma_z\rangle_{t=0} \\ -i\langle\sigma_y\rangle_{t=0} \end{pmatrix} e^{-iz_\sigma t} \quad , \\ - \,2\alpha \, \frac{1}{(\tilde{\Delta}t)^2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \langle\sigma_z\rangle_{t=0} \\ -i\langle\sigma_y\rangle_{t=0} \end{pmatrix} e^{-iz_0 t} \quad . \tag{156}$$

In this result all logarithmic terms at high energies $\sim (\alpha \ln \frac{D}{\tilde{\Delta}})^k$ have been resummed in the renormalized tunneling $\tilde{\Delta}$, and all logarithmic terms at low energies (or large times) $\sim (\alpha \ln(\tilde{\Delta} t))^k$ are contained in $s(t)$. For the pre-exponential function $F_0^b(t)$ it turns out that, in leading order, no logarithmic terms are present at large times. This has to be contrasted to the solution within the noninteracting blip approximation (NIBA) [2], where, for $\langle \sigma_z \rangle_{t=0} = 1$ and $\langle \sigma_y \rangle_{t=0} = 0$, one obtains

$$\langle \sigma_z \rangle(t)_{NIBA} \;=\; e^{-\frac{\Gamma}{2}t} \cos(\tilde{\Delta} t) \;-\; 2\alpha \, \frac{1}{(\tilde{\Delta} t)^{2-2\alpha}} \,, \tag{157}$$

whereas the correct result from (156) reads

$$\langle \sigma_z \rangle(t) \;=\; e^{-\frac{\Gamma}{2}t} \cos(\tilde{\Delta} t) \;-\; 2\alpha \, \frac{1}{(\tilde{\Delta} t)^{2}} \, e^{-\Gamma t} \,. \tag{158}$$

Besides the missing exponential part in the second term, which has already been discussed at the end of Section 4, the NIBA predicts a different exponent for the pre-exponential power law. This shows that power-law exponents of pre-exponential functions can only be calculated by resumming consistently all logarithmic terms for long times. The E-RTRG method predicts that no such logarithmic terms are present for $\langle \sigma_{y,z} \rangle(t)$ but they appear for $\langle \sigma_x \rangle(t)$ within the logarithmic function $s(t)$. The leading power-law behaviour $\sim (\frac{1}{t})^2$ of the pre-exponential function is the same as for the Kondo model and can also be obtained from perturbative calculations [29]. There are always two terms with different decay rates Γ and $\Gamma/2$ for the time evolution. If one transforms to the exact eigenbasis $\underline{e}_{1/2} = \frac{1}{\sqrt{2}}(|+\rangle \pm |-\rangle)$ of the local system, the expectation values $\langle \gamma_i \rangle$ of the Pauli matrices in the new basis are related to the ones of the original basis by $\langle \gamma_x \rangle = \langle \sigma_z \rangle$, $\langle \gamma_y \rangle = -\langle \sigma_y \rangle$ and $\langle \gamma_z \rangle = \langle \sigma_x \rangle$. Thus the Markovian term $\sim e^{-\Gamma t}$ from the pole contribution describes the decay of the diagonal matrix elements of the density matrix in the new basis, whereas the one $\sim e^{-(\Gamma/2)t} e^{\pm \tilde{\Delta} t}$ corresponds to the decay of the nondiagonal matrix elements. Therefore, Γ is called the relaxation rate, whereas $\Gamma/2$ is the decoherence rate, in accordance with the general rule that, in the absence of pure dephasing, the relaxation rate is always twice as large as the decoherence rate.

For large energies $|E| \gg \tilde{\Delta}$, one needs the function $Z'(E)$ to determine the regime of **short times** $t \ll \frac{1}{\tilde{\Delta}}$ from (42) (the contribution from the exponential is a small correction and can be neglected). One obtains the result

$$Z'(E) \;=\; \sum_{\sigma=\pm} \begin{pmatrix} 1 & 0 \\ 0 & Z_\sigma(E) \end{pmatrix} \otimes \tau_\sigma \quad , \quad Z_\pm(E) \;\approx\; \left(\frac{-iE}{D} \right)^{2\alpha} \tag{159}$$

This gives rise to the universal short time behavior

$$\rho(t) \;=\; \begin{pmatrix} 1 & 0 \\ 0 & (\frac{1}{Dt})^{2\alpha} \end{pmatrix} \otimes 1 \, \rho_{t=0} \quad , \tag{160}$$

or

$$\langle \sigma_{x,y} \rangle(t) \;=\; (\frac{1}{Dt})^{2\alpha} \, \langle \sigma_{x,y} \rangle_{t=0} \quad , \quad \langle \sigma_z \rangle(t) \;=\; \langle \sigma_z \rangle_{t=0} \quad . \tag{161}$$

This agrees with previous predictions and can also be obtained from the exact solution (20) at $\Delta = T = 0$ in the universal regime $t \gg \frac{1}{D}$. Again we can see that all logarithmic terms $\sim (\alpha \ln(Dt))^k$ have been resummed in this result.

6.3 Interacting resonant level model

Finally we discuss the IRLM for the special case of a single reservoir with chemical potential $\mu = 0$ and zero level position $\epsilon = 0$ (i.e. in resonance with the reservoir). As discussed in Section 2, this model can be mapped to the ohmic spin boson model close to the exactly solvable point $\alpha = \frac{1}{2}$. In particular, we want to understand where the crossover from coherent to incoherent time evolution by changing the sign of $U = 1 - \sqrt{2\alpha}$ comes from. We follow Ref. [8, 9], where the IRLM has been studied by using E-RTRG and functional RG.

We concentrate on the time evolution of the occupation $\langle n \rangle(t)$ of the local level which is related via Eq. (22) to the expectation $\langle \sigma_z \rangle(t)$ within the spin boson model by $2\langle n \rangle(t) - 1 = \langle \sigma_z \rangle(t)$. For $\langle \sigma_z \rangle(t)$, one can show that the result can be written in the form $\langle \sigma_z \rangle(t) = P(t) \langle \sigma_z \rangle_{t=0}$, with

$$P(t) \;=\; \frac{i}{2\pi} \int_{-\infty+i0^+}^{\infty+i0^+} dE\, e^{-iEt}\, \frac{1}{E \,+\, i\Gamma_1(E)} \quad . \tag{162}$$

For this special case there is no Z'-factor and $\Gamma_1(E)$ is a slowly varying logarithmic function describing the energy dependent charge relaxation rate. It is determined from the RG equations

$$\frac{\partial}{\partial E}\Gamma_1(E) \;=\; -g\, R_2(E)\, \Gamma_1(E) \quad , \quad \frac{\partial}{\partial E}\Gamma_2(E) \;=\; -g\, R_1(E)\, \Gamma_1(E) \quad , \tag{163}$$

where $g = 2U - U^2 = 1 - 2\alpha$ and the resolvents $R_{1/2}(E)$ are defined by

$$R_1(E) \;=\; \frac{1}{E \,+\, i\Gamma_1(E)} \quad , \quad R_2(E) \;=\; \frac{1}{E \,+\, i\Gamma_2(E)/2} \quad . \tag{164}$$

The initial conditions are given by $\Gamma_{1/2}(E = iD) = \Gamma^{(0)}$. $\Gamma_2(E)/2$ is also a slowly varying logarithmic function and describes the energy dependent broadening of the local level corresponding to the decoherence mode for nondiagonal matrix elements of the local density matrix w.r.t. the charge states (note, however, that such elements can not be prepared). As we will see below the subtle coupling of the two RG equations for $\Gamma_{1/2}(E)$ leads to the interesting effect that, for $g > 0$, the resolvent $R_1(E)$ can have poles with a finite real part although the local system has no finite excitation energy.

We start by solving the RG equations at **high energies** $E \gg \Gamma_{1/2}(E)$. Neglecting $\Gamma_{1/2}(E)$ on the r.h.s. of the RG equations, we find the solution

$$\Gamma_{1/2}(E) \;=\; \Gamma^{(0)} \left(\frac{D}{-iE}\right)^g \;=\; \tilde{\Delta} \left(\frac{\tilde{\Delta}}{-iE}\right)^g , \tag{165}$$

where

$$\tilde{\Delta} \;=\; \Gamma^{(0)} \left(\frac{D}{\tilde{\Delta}}\right)^g \;=\; \Gamma^{(0)} \left(\frac{D}{\Gamma^{(0)}}\right)^{g/(1+g)} \tag{166}$$

is the renormalized tunneling which is kept fixed in the scaling limit $D \to \infty$ and $\alpha, \Gamma^{(0)} \to 0$. Using the relation $g = 1 - 2\alpha$ and $\Gamma^{(0)} = \frac{\Delta^2}{D}$ to the spin boson model, one can see that it is identical to the definition (149) of the renormalized tunneling for the spin boson model. As discussed in detail in Refs. [7, 9], the solution at high energies contains all leading logarithmic

terms $\sim (U \ln \frac{D}{-iE})^k$ and all subleading ones $\sim U(U \ln \frac{D}{-iE})^k$. From the solution (166) at high energies we can calculate with (42) the time evolution for **short times** $t \ll 1/\tilde{\Delta}$ as

$$P(t) \approx e^{-\Gamma_1(1/t)t} \approx e^{-(\tilde{\Delta}t)^{1+g}}, \tag{167}$$

i.e. the relaxation rate in the exponent is cut off at the energy scale $1/t$. In contrast to the spin boson model at small α and the Kondo model, there is no Z'-factor and therefore the exponential provides the leading order. Expanding the exponential we find $P(t) = 1 - (\tilde{\Delta}t)^{1+g}$ in agreement with previous results [2]. Since $(\tilde{\Delta}t)^{1+g} = (Dt)^g \Gamma^{(0)} t$, we see again that all logarithmic terms $\sim (g \ln(Dt))^k$ have been resummed for small times.

Next we study the analytic structure of the resolvent $R_1(E)$ to find the time evolution for intermediate and long times. As we will show below, for positive $g > 0$, $R_1(E)$ has two poles at $z_\pm$ (followed by a branch cut with jump of $O(g^2)$ which can be neglected) and one branch cut starting at z_0 (with jump of $O(g)$), where the singularities z_n, $n = 0, \pm$, are given by

$$z_0 = -i\frac{\tilde{\Delta}}{2} \quad , \quad z_\pm = \pm\Omega - i\tilde{\Delta} \quad , \quad \Omega = \pi g \tilde{\Delta} \quad . \tag{168}$$

For $g < 0$, there is only a branch cut starting at z_0. Thereby, z_0 is the position of the pole of the resolvent $R_2(E)$, i.e. z_0 and $z_\pm$ can be determined from the equations

$$z_\pm + i\Gamma_1(z_\pm) = 0 \quad , \quad z_0 + i\Gamma_2(z_0)/2 = 0 \quad . \tag{169}$$

Note that, in contrast to the singularities (148) for the spin boson model at small α, for the IRLM (or the spin boson model at $\alpha = \frac{1}{2}$) the renormalized tunneling determines the rate and not the oscillation frequency. Furthermore, we note that the pole of $R_1(E)$ describes the charge relaxation mode, whereas for the spin boson model at small α it corresponds to the decoherence mode w.r.t. the exact eigenstates of the local system. Therefore, z_0 corresponds to the decoherence mode for the IRLM and its imaginary part is half of the one of the relaxation poles $z_\pm$. To derive the result for the positions of the singularities we solve the RG equations for intermediate and small energies $|E| \lesssim \tilde{\Delta}$ but $g \ln \frac{\tilde{\Delta}}{|E - z_n|} \ll 1$, i.e. E should not be exponentially close to the singularities. Expanding in the small parameter $g \ln \frac{\tilde{\Delta}}{|E-z_n|} \ll 1$ and fixing the integration constants by comparing with the solution (165) at high energies in the usual way, we find

$$\Gamma_1(E)/\tilde{\Delta} \approx 1 - g \ln \frac{-iE + \Gamma_2(E)/2}{\tilde{\Delta}} \ , \ \Gamma_2(E)/\tilde{\Delta} \approx 1 - g \ln \frac{-iE + \Gamma_1(E)}{\tilde{\Delta}}. \tag{170}$$

In contrast to the corresponding equation (142) for the Kondo model, there is a subtle coupling of the singularites of $\Gamma_1(E)$ and $\Gamma_2(E)$, which leads to the new feature that $z_\pm$ obtains a finite real part for $g > 0$. We note that although the equations can not be used for E exponentially close to the singularities, they can be used for $|E - z_n| \sim g^2$ since $g \ln(g) \ll 1$ for $g \ll 1$. Therefore, the equations can be used to determine the positions of the branching points of $\Gamma_{1/2}(E)$ up to $O(g)$. From the equations we can see that $\Gamma_1(E)$ ($\Gamma_2(E)$) have a branch cut with jump of $O(g)$ starting at the branching point of the logarithmic function where (169) is fulfilled, i.e. at z_0 ($z_\pm$). Thereby, the branch cut of $\Gamma_2(E)$ starting at $z_\pm$ leads also to a branch cut for $\Gamma_1(E)$ at the same position but this branch cut has a jump of $O(g^2)$ and can be neglected.

Inserting the leading order results $\Gamma_{1/2}(E) \approx \tilde{\Delta}$, $z_0 \approx -i\tilde{\Delta}/2$ and $z_\pm \approx -i\tilde{\Delta}$ on the r.h.s. of (170), we find for the position of the singularities the result (168)

$$2iz_0/\tilde{\Delta} = \Gamma_2(z_0)/\tilde{\Delta} \approx 1 - g\ln(-iz_0/\tilde{\Delta}+1) \approx 1 - g\ln\left(-\frac{1}{2}+1\right) \approx 1 \quad , \tag{171}$$

$$\begin{aligned} iz_\pm/\tilde{\Delta} = \Gamma_1(z_\pm)/\tilde{\Delta} &\approx 1 - g\ln\left(-iz_\pm/\tilde{\Delta}+\frac{1}{2}\right) \approx 1 - g\ln\left(-1 \mp i\Omega/\tilde{\Delta}+\frac{1}{2}\right) \\ &\approx 1 \pm i\pi g \quad . \end{aligned} \tag{172}$$

Due to the analytic structure of the resolvent $R_1(E)$ the time evolution can be written as

$$P(t) \;=\; \theta(g) \sum_{\sigma=\pm} e^{-iz_\sigma t} \;+\; F_0^b(t)\, e^{-iz_0 t} \;=\; \theta(g)\, 2\, \cos(\Omega t)\, e^{-\tilde{\Delta}t} \;+\; F_0^b(t)\, e^{-(\tilde{\Delta}/2)t}\,, \tag{173}$$

where the first term involves the contribution from the isolated poles (we have neglected corrections of $O(g)$ to the residuum) and the second term involves the analog of the branch cut integral (48), which can be written as

$$F_0^b(t) \;=\; \frac{1}{\pi}\,\mathrm{Im} \int_0^\infty dx\, e^{-xt}\, \frac{1}{-i(z_0 \,+\, i\Gamma_1(z_0 - i/t + 0^+)) \,-\, x} \quad . \tag{174}$$

For **intermediate and long times** $t \gtrsim \frac{1}{\tilde{\Delta}}$ but $g\ln(\tilde{\Delta}t) \ll 1$, $F_0^b(t)$ can be evaluated by using the result (170), where we obtain $-i\Gamma_1(z_0 - i/t + 0^+) = z_+(1 + O(g\ln(\tilde{\Delta}t))$. In particular one has to consider the fact that $x \sim \frac{1}{t}$ can not be neglected compared to the difference $|z_0 - z_\pm| \sim \tilde{\Delta}$ for intermediate times $t \sim \frac{1}{\tilde{\Delta}}$. This time regime is of particular interest here since the exponentials of the time evolution (173) decay on the time scale $\frac{1}{\tilde{\Delta}}$. Therefore, the integral (174) has to be calculated more carefully in terms of the exponential integral $E_1(z)$

$$F_0^b(t) \;=\; -\frac{1}{\pi}\,\mathrm{Im}\,\left\{e^{-i(z_0 - z_+)t}\, E_1(-i(z_0 - z_+)t)\right\} \quad . \tag{175}$$

This result has been used in Refs. [8, 9] to discuss the competition between the oscillating (i.e. coherent) and the purely decaying (i.e. incoherent) term of the time evolution in Eq. (173). Since the incoherent term decays on a longer time scale it turns out that it wins very rapidly such that the coherent term leads only to a few number of oscillations, in contrast to the physics of a classical damped harmonic oscillator. For **long times** $t \gg \frac{1}{\tilde{\Delta}}$ but still $g\ln(\tilde{\Delta}t) \ll 1$, the incoherent term dominates and, using the asymptotic expansion $E_1(z) = e^z/z$ of the exponential integral, one obtains

$$F_0^b(t) \;\approx\; -4g\,\frac{1}{\tilde{\Delta}t} \quad , \tag{176}$$

i.e. a power law $\sim 1/t$ typical for models with charge fluctuations.

Finally, for **exponentially large times** $g\ln(\tilde{\Delta}t) \;\sim\; O(1)$, we need the solution for $\Gamma_1(E)$ for energies E exponentially close to the branching point z_0. In this regime, we can replace $\Gamma_2(E)/2 \to iz_0$ on the r.h.s. of the RG equation (163) for $\Gamma_1(E)$, which gives the solution

$$\Gamma_1(E) \;=\; \tilde{\Delta}\,\left(\frac{\tilde{\Delta}}{-i(E - z_0)}\right)^g \quad , \tag{177}$$

where the integration constant has been fixed by comparison with the solution (170) at intermediate and small energies. Using this solution for the evaluation of the branch cut integral (174) for exponentially large times, we can neglect x in the denominator and find with $\Gamma_1(z_0 - i/t + 0^+) = \tilde{\Delta}(\tilde{\Delta}t)^g + 2\pi i g \tilde{\Delta}(\tilde{\Delta}t)^g$ the result (neglecting terms of $O(g^2)$ in the denominator)

$$F_0^b(t) \approx -g \frac{1}{(1/2 - (\tilde{\Delta}t)^g)^2} \frac{1}{(\tilde{\Delta}t)^{1-g}} \quad . \tag{178}$$

This result holds for all times $t \gg \frac{1}{\tilde{\Delta}}$. For long times with $g \ln(\tilde{\Delta}t) \ll 1$ it reduces to the result (176). However, for exponentially large times where $(\tilde{\Delta}t)^g$ is some number of $O(1)$, the result changes. In the extreme regime $(\tilde{\Delta}t)^{|g|} \gg 1$, it reduces to

$$F_0^b(t) \approx -g\,[1 + 3\theta(-g)] \frac{1}{(\tilde{\Delta}t)^{1+|g|}} \quad . \tag{179}$$

This result agrees with the prediction of the NIBA [2] and its improved version [24] (where the exponential term $e^{-(\tilde{\Delta}/2)t}$ has also been obtained, see Eq. (173)). However, as we have seen, it holds only for extremely long times and, for $g > 0$, the prefactor is different from the result (176) for more realistically long times $\tilde{\Delta}t \gg 1$ with $g \ln(\tilde{\Delta}t) \ll 1$. Therefore, we see that the regime of long times is very subtle and the result can change significantly by entering the regime of exponentially large times.

Finally, as already mentioned in Section 5, it has not yet been studied to a full extent how the RG equation (163) looks like in higher orders in the tunneling. There is some evidence that all higher order terms in Γ are of the form

$$U^n \left(\frac{\Gamma_i}{\tilde{\Delta}}\right)^k (\ln \frac{E - z_n}{\tilde{\Delta}})^l \quad \text{for} \quad n = 1, 2 \quad , \quad U^n \frac{\Gamma_i}{E - z_n} \left(\frac{\Gamma_i}{\tilde{\Delta}}\right)^k \text{for} \quad n > 2 \quad , \tag{180}$$

i.e., after integration, either vanish in the limit $E \to z_n$ or contribute to higher orders in U, but this is still under investigation. Furthermore, the results have been compared to functional RG in Refs. [8, 9], where all orders in the tunneling have been resummed keeping only the lowest order term in the Coulomb interaction. The numerical results of functional RG agree quite nicely with the analytical result (175) for intermediate and long times and, in particular for extremely long times, the result (179) has been confirmed analytically by functional RG. Therefore, there is good evidence that also within E-RTRG higher orders in the tunneling will not change the results at least in leading order in U.

Acknowledgments. I am particularly thankful to M. Wegewijs for a thorough reading of the manuscript.

References

[1] H. Mori, Progress of Theoretical Physics **33**, 423 (1965); R. Zwanzig and R. D. Mountain, Journal of Chemical Physics **43**, 4464 (1965); H. Grabert, *Projection Operator Techniques in Nonequilibrium Statistical Mechanics*, 1st ed., Springer Tracts in Modern Physics, Vol. 95 (Springer, Berlin, 1982) p. 166; E. Fick and G. Sauermann, *The Quantum Statistics of Dynamic Processes* (Springer Series in Solid-State Sciences 86, Springer-Verlag, Berlin, 1990); C. W. Gardiner, *Quantum Noise* (Springer Series in Synergetics 56, Springer-Verlag, Berlin, 1991); R. Zwanzig, *Nonequilibrium Statistical Mechanics*, 3rd ed. (Oxford University Press, New York, 2001) p. 240.

[2] A. J. Leggett *et al*., Rev. Mod. Phys. **59**, 1 (1987); H. Grabert, P. Schramm, and G. L. Ingold, Phys. Rep. **168**, 115 (1988); U. Weiss, *Quantum Dissipative Systems* (World Scientific Publishing Company, Singapore, 2012).

[3] H. Schoeller, Eur. Phys. J. Spec. Top. **168**, 179 (2009).

[4] H. Schoeller and F. Reininghaus, Phys. Rev. B **80**, 045117 (2009); *ibid*. Phys. Rev. B **80**, 209901(E) (2009).

[5] M. Pletyukhov, D. Schuricht, and H. Schoeller, Phys. Rev. Lett. **104**, 106801 (2010).

[6] M. Pletyukhov and H. Schoeller, Phys. Rev. Lett. **108**, 260601 (2012); F. Reininghaus, M. Pletyukhov, and H. Schoeller, in preparation.

[7] C. Karrasch *et al*., Europhys. Lett. **90**, 30003 (2010); S. Andergassen *et al*., Phys. Rev. B **83**, 205103 (2011).

[8] D. M. Kennes *et al*., Phys. Rev. Lett. **110**, 100405 (2013).

[9] O. Kashuba *et al*., Phys. Rev. B **88**, 165133 (2013).

[10] O. Kashuba and H. Schoeller, Phys. Rev. B **87**, 201402(R) (2013).

[11] S. Kehrein, Phys. Rev. Lett. **95**, 056602 (2005). D. Lobaskin and S. Kehrein, Phys. Rev. B **71**, 193303 (2005).

[12] R. Gezzi, T. Pruschke, and V. Meden, Phys. Rev. B **75**, 045324 (2007); S. G. Jakobs, V. Meden, and H. Schoeller, Phys. Rev. Lett. **99**, 150603 (2007); W. Metzner *et al*., Rev. Mod. Phys. **84**, 299 (2012); D. M. Kennes *et al*., Phys. Rev. B **85**, 085113 (2012).

[13] F. B. Anders and A. Schiller, Phys. Rev. Lett. **95**, 196801 (2005); F. B. Anders, R. Bulla, and M. Vojta, Phys. Rev. Lett. **98**, 210402 (2007); A. Hackl *et al*., Phys. Rev. Lett. **102**, 219902 (2009).

[14] A. Daley *et al*., J. Stat. Mech.: Theor. Exp. P04005 (2004); S. R. White and A. Feiguin, Phys. Rev. Lett. **93**, 076401 (2004); P. Schmitteckert, Phys. Rev. B **70**, 121302 (2004); F. Heidrich-Meisner, A. E. Feiguin, and E. Dagotto, Phys. Rev. B **79**, 235336 (2009).

[15] S. Weiss *et al*., Phys. Rev. B **77**, 195316 (2008).

[16] T. L. Schmidt *et al*., Phys. Rev. B **78**, 235110 (2008).

[17] F. Lesage and H. Saleur, Phys. Rev. Lett. **80**, 4370 (1998); A. Schiller and S. Hershfield, Phys. Rev. B **62**, R16271 (2000); A. Komnik, Phys. Rev. B **79**, 245102 (2009).

[18] A. Rosch, J. Kroha and P. Wölfle, Phys. Rev. Lett. **87**, 156802 (2001); A. Rosch *et al*., Phys. Rev. Lett. **90**, 076804 (2003); L. I. Glazman and M. Pustilnik, in *Nanophysics: Coherence and Transport* (H. Bouchiat et al., Elsevier, 2005), p. 427; T.Korb *et al*., Phys. Rev. B **76**, 165316 (2007);

[19] R. B. Saptsov and M. R. Wegewijs, Phys. Rev. B **86**, 235432 (2012).

[20] R. B. Saptsov and M. R. Wegewijs, arXiv:1311.1368.

[21] D. Schuricht and H. Schoeller, Phys. Rev. B **80**, 075120 (2009); S. Y. Mueller *et al*., arXiv:1211.7072.

[22] O. Kashuba, H. Schoeller, and J. Splettstoesser, EPL **98** (2012) 57003.

[23] A. A. Slutskin *et al*., Europhys. Lett. **96**, 40003 (2011).

[24] R. Egger, H. Grabert, and U. Weiss, Phys. Rev. E **55**, R3809 (1997).

[25] P. W. Anderson, J. Phys. C **3**, 2436 (1970); A. C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge University Press, 1997).

[26] S. Goettel, F. Reininghaus, and H. Schoeller, in preparation.

[27] A. Hackl *et al*., Phys. Rev. Lett. **102**, 219902 (2009); A. Hackl, M. Vojta, and S. Kehrein, Phys. Rev. B **80**, 195117 (2009).

[28] L. D. Contreras-Pulido *et al*., Phys. Rev. B **85**, 075301 (2012).

[29] D. P. DiVincenzo and D. Loss, Phys. Rev. B **71**, 035318 (2005).

B 6 Edge Magnetism in Graphene [1]

M. Schmidt
Institut für Theoretische Festkörperphysik
RWTH Aachen

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Graphene is a two-dimensional material consisting of a honeycomb lattice of carbon atoms [1]. It is due to this lattice geometry that the relevant part of the band structure (the part close to the Fermi level) is well approximated by the famous Dirac cones (one cone for each K point in the Brillouin zone). The effects of electron-electron interactions in graphene are, as will be explained in more detail below, suppressed for usual bulk graphene samples. However, we will see in this chapter that this suppression only works as long as the honeycomb lattice is perfect. For real-life graphene samples with edges and defects, electron-electron interactions come back into the game and are responsible for a rather unusual phenomenon: edge magnetism. But some might ask, why one would consider edge magnetism to be more unusual than the standard types of ferro- and antiferromagnetism. In short, this is due to the low dimensionality of the phenomenon (edges are one-dimensional). Mean-field approximations (with DFT being much like a mean-field approximation in this context) are dangerous in one dimension and we will see that it is important to go beyond this approximation in order to describe the true ground state and its qualitative features correctly.

But before we discuss the details of edge magnetism, let us first understand how we describe electronic interactions and why they are not of much importance in the bulk of graphene. Thereby we will also introduce the very basic description of interacting graphene that we will use throughout this chapter. As far as the non-interacting physics is concerned, the tight-binding Hamiltonian in second quantization

$$\hat{H}_0 = -\sum_{i,j,\sigma} J_{ij} \hat{c}^\dagger_{i\sigma} \hat{c}_{j\sigma} \tag{1}$$

with $\sigma = \uparrow, \downarrow$ the z-direction of the electron spin, J_{ij} the hopping amplitude from site i to site j, and $\hat{c}_{i\sigma}$ ($\hat{c}^\dagger_{i\sigma}$) the 2nd quantized annihilation (creation) operator of an electron with spin σ in the π orbital at site i, is an excellent approximation for graphene. For our purposes it is even sufficient to approximate the hopping $J_{ij} = J \approx 3$ eV for i, j being nearest neighbors and $J_{ij} = 0$ otherwise.

We have formulated the non-interacting electrons in the π orbital basis. The same can be done also for the electron-electron interaction. A reasonable starting point for the Coulomb repulsion of the electrons is the general Hamiltonian

$$\hat{H}_{\rm int} = U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} + \sum_{i<j} V_{ij} \hat{n}_i \hat{n}_j, \tag{2}$$

where $\hat{n}_{i\sigma} = \hat{c}^\dagger_{i\sigma} \hat{c}_{i\sigma}$ is the occupation number of the π orbital at site i with spin σ and $\hat{n}_i = \hat{n}_{i\uparrow} + \hat{n}_{i\downarrow}$. U is the on-site Hubbard energy, which has to be be paid if two electrons occupy the same site. $V_{ij} \propto 1/|\mathbf{r}_i - \mathbf{r}_j|$ is the long-range Coulomb repulsion. For the considerations we aim at, the Hubbard part of $\hat{H}_{\rm int}$ is most important. We thus set $V_{ij} = 0$ in the remainder of this chapter and keep only the Hubbard term.[2]

We now argue that U is particularly large in graphene, compared to other effectively two-dimensional electron systems. Given a Wannier wave function $\phi^{\rm W}_i(\mathbf{r})$ for the electron at the

[2] This should not be interpreted to the effect that the long-range part of the interaction is completely irrelevant. In fact, it has some pronounced and experimentally observed effects on the electronic spectrum. Also for certain parameters it might drive the system into a quantum spin-Hall state. The point here is, however, that *the long-range part is irrelevant for the magnetic properties.*

ith site, such that $c_{i\sigma}^\dagger = \int \mathrm{d}^3\mathbf{r}\phi_i^{\mathrm{W}}(\mathbf{r})\hat{\psi}_\sigma(\mathbf{r})$, where $\hat{\psi}_\sigma(\mathbf{r})$ is the 2nd quantized field operator with spin σ, the strength of the Hubbard interaction U can be estimated by

$$U \approx e^2 \int \mathrm{d}^3\mathbf{r} \int \mathrm{d}^3\mathbf{r}' \frac{|\phi_i^{\mathrm{W}}(\mathbf{r})|^2 |\phi_i^{\mathrm{W}}(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|}. \tag{3}$$

Obviously, the more localized the Wannier functions are, the larger is U. In a typical two-dimensional electron gas, the Wannier functions are extended over several lattice constants in the direction perpendicular to the 2D plane. This is because the confinement potential, which is generated by building layered structures from semiconductors with slightly different band structures, is relatively shallow and varies slowly. The confinement potential of the Wannier functions in graphene is the atomic potential of the carbon atoms, and is much larger. Therefore, this leads to a much stronger confinement. In other words: in graphene the electrons are squeezed together in a two-dimensional sheet, the thickness of which is of the order of an Angstrom. Thus, the energy U which must be paid for two electrons sitting at the same place in the other two dimensions is considerably larger than for conventional two-dimensional electron gasses.

From the point of view of electron-electron interactions, one of the most important features of the Dirac-cone band structure is the linearly vanishing density of states near the charge neutrality point. A small density of states at a certain energy means (in a sense) that "there are not many electrons in this energy range to which something could happen". But for the electron-electron interaction to actually have an impact, electrons are obviously needed. A slightly less colloquial formulation of this thought is provided by the Stoner criterion $D(\epsilon_F)U > 1$, which states that the Hubbard interaction U can drive a phase transition only if it is stronger than the inverse density of states at the Fermi level $D(\epsilon_F)$. However, since the density of states is very small or even vanishes in graphene, extremely large U would be needed for driving a phase transition (in graphene this would be the transition to a bulk antiferromagnetic phase). Also elaborate techniques, such as quantum Monte-Carlo simulations (see also Chapter B4) show that U is not large enough for generating antiferromagnetism in graphene. And of course we know from experiment that the graphene bulk is not antiferromagnetic. But now we understand that it is the vanishing density of states, a feature of the Dirac-cone band structure, which is responsible for the suppression of interaction effects. The interaction itself, measured in absolute values (U) is not small at all.

So we are faced with the following situation: Due to the strong confinement to two dimensions, the Coulomb repulsion in graphene is very large in absolute terms. But the vanishing density of states at the charge neutrality point suppresses the effect of the electron-electron interactions. Furthermore, the vanishing density of states is a result of the perfect honeycomb lattice of graphene. But what happens if the lattice is not perfect so that we cannot count on a small density of states? This is the setting which will be discussed in this chapter.

2 The zigzag edge – edge states

In this section, the electron spin is not important. Thus, we drop the label σ for all electron operators. It can be reintroduced at the end by the usual doubling of the theory, i.e., the total Hamiltonian is a sum of a spin-up and a spin-down Hamiltonian. Consider the hopping Hamiltonian for graphene $\hat{H}_0$. If we introduce explicit site labelling $i = (n_1 n_2 s)$ with $n_1\mathbf{a}_1 + n_2\mathbf{a}_2$ the

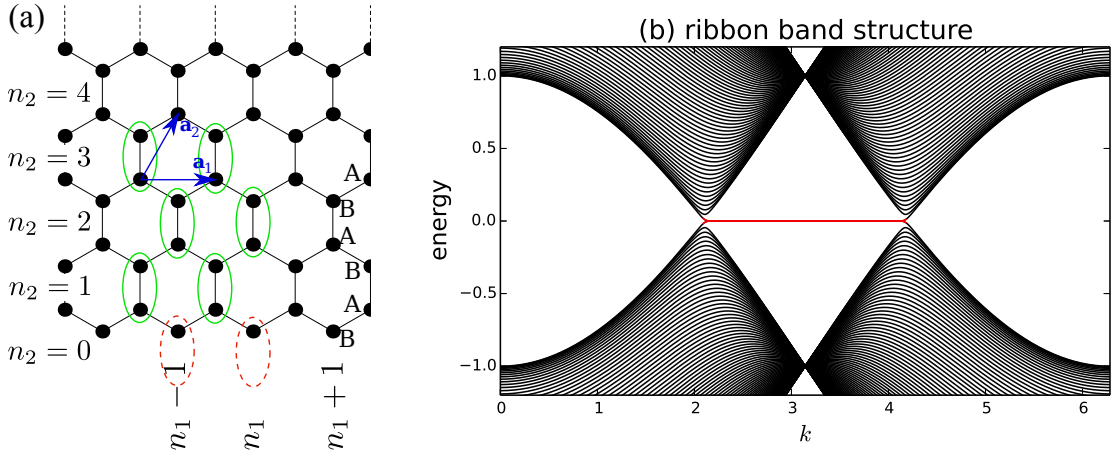


Fig. 1: *(a) The definitions of the honeycomb lattice.* $\mathbf{a}_1, \mathbf{a}_2$ *are the Bravais lattice vectors. The green circles are the unit cells with sublattice sites A and B. The red-dashed circles are the truncated unit cells at the zigzag edge. The coordinates* n_1 *and* n_2 *are indicated. (b) The band structure of a graphene ribbon with* $N_2 = 100$*. The black curves, forming cone-like structures, are the bulk bands. The bands with nearly zero energy in red are the edge states. The Fermi level is at or at least close to zero energy.*

vector of the unit cell and $s = A, B$ the sublattice index (see Fig. 1a), the Hamiltonian reads

$$\hat{H}_0 = -J \sum_{\langle i,j \rangle} \hat{c}_i^\dagger \hat{c}_j = -J \sum_{n_1 n_2} \hat{c}_{n_1 n_2 A}^\dagger (\hat{c}_{n_1 n_2 B} + \hat{c}_{n_1 n_2 - 1 B} + \hat{c}_{n_1 + 1 n_2 - 1 B}) + H.c. \tag{4}$$

The particular representations of $\mathbf{a}_1, \mathbf{a}_2$ are not important for us. We assume that they are chosen such that they generate a honeycomb lattice. All that matters is already stated in the Hamiltonian.

Next we introduce a zigzag edge along the $\mathbf{a}_1$ direction. Thus, the crystal momentum along the $\mathbf{a}_1$ direction is a good quantum number – a fact we make use of by performing a partial Fourier transformation to the momentum along the edge (k)

$$\hat{c}_{n_1 n_2 s} = \frac{1}{\sqrt{N_1}} \sum_k e^{ikn_1} \hat{c}_{k n_2 s}, \tag{5}$$

with N_1 the number of unit cells in $\mathbf{a}_1$ direction. The transformed Hamiltonian then takes the form of a one-dimensional chain with k-dependent alternating hopping (see Fig. 2a)

$$\hat{H}_0 = -J \sum_k \sum_{n_2=1}^{\infty} \left[\hat{c}_{k n_2 A}^\dagger \hat{c}_{k n_2 B} + \hat{c}_{k n_2 A}^\dagger (1 + e^{ik}) \hat{c}_{k n_2 - 1 B} + H.c. \right], \tag{6}$$

where we have restricted the sum to $n_2 \geq 1$ such that the edge in Fig. 1a is described.

If we further restrict $n_2 \leq N_2$, this Hamiltonian can be easily diagonalized numerically. The spectrum as a function of the momentum k yields the band structure of graphene with an edge. It is shown in Fig. 1b. Instead of a two-dimensional band structure we obtain many one-dimensional bands. Technically this is because we only have one translationally invariant direction due to the edge. But one might expect that this change in boundary conditions has the sole effect of performing a zone folding of the full two-dimensional band structure. In Fig. 1b we

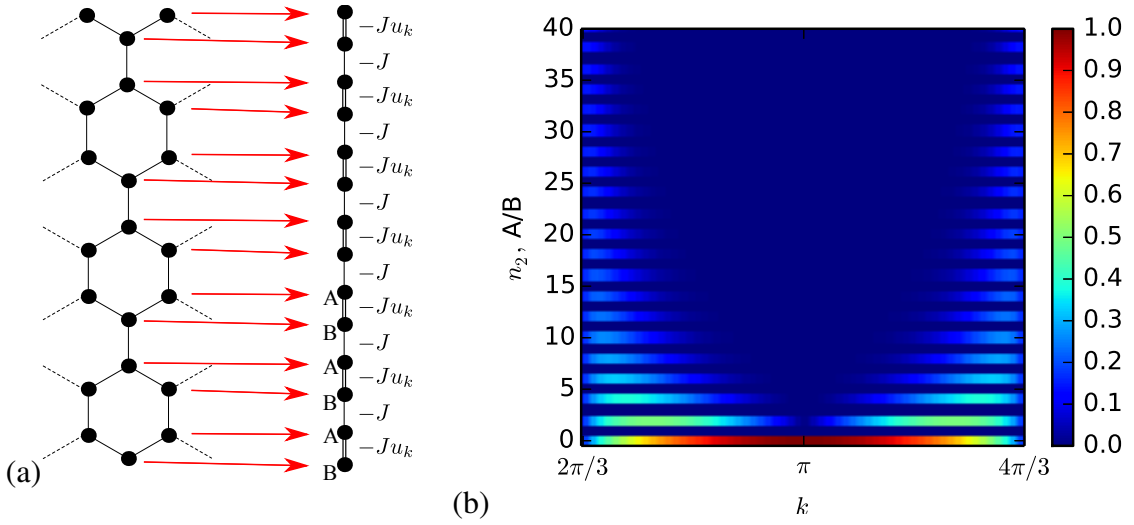


Fig. 2: *(a) The effective one-dimensional chain with k-dependent hoppings $-J$ and $-Ju_k$, with $u_k = 1 + e^{ik}$. (b) The transverse edge state wave function $|\phi_k(n_2)| = \mathcal{N}_k|u_k|^{n_2}$. Only the B sublattice sites have non-zero weight.*

see that this is nearly true, indeed. Most bands are the sections of the 2D Dirac cones. But there is an additional band (in red) at practically zero energy which appears to connect the two Dirac cones [2]. This additional band cannot be obtained by zone folding, but we will now calculate the wave functions of these zero energy states directly. One may expect already at this point that they will not be constructed from plain waves as the bulk states which form the Dirac cones.
We note that these wave functions have nearly zero energy. Closer (numerical) inspection of the spectrum shows that the energy is in fact exponentially small in the width N_2 of the ribbon. Thus, we expect that for $N_2 \to \infty$ there should be zero-energy eigenstates of the Hamiltonian $\hat{H}_0$ in Eq. (6). These we would like to find with the Ansatz

$$\hat{e}^\dagger_{kB} = \sum_{n_2=0}^{\infty} \phi_{kB}(n_2)\hat{c}^\dagger_{kn_2B} \tag{7}$$

for the edge state creation operator, with $\phi_{kB}(n_2) \in \mathbb{C}$ the wave function of the edge state. Note that we have already chosen our Ansatz such that the wave function is supported only on the B sublattice. A more general Ansatz leads to the same result, of course. The general commutator relation $[\hat{H}_0, \hat{e}^\dagger_{kB}] = \epsilon_k \hat{e}^\dagger_{kB}$ for eigenstates $\hat{e}^\dagger_{kB}$ with energy ϵ reduces in our case $\epsilon = 0$ to the notion that $\hat{e}^\dagger_{kB}$ commutes with $\hat{H}_0$. We find

$$[\hat{H}_0, \hat{e}^\dagger_{kB}] = -J \sum_{n_2=1}^{\infty} c^\dagger_{kn_2A}[\phi_{kB}(n_2) + \underbrace{(1 + e^{ik})}_{=u_k} \phi_{kB}(n_2 - 1)] \stackrel{!}{=} 0, \tag{8}$$

which is only zero if $\phi_{kB}(n_2) = -u_k\phi_{kB}(n_2 - 1)$. This relation can only be fulfilled for $\phi_{kB}(n_2) \propto (-u_k)^{n_2}$. The constant of proportionality $\mathcal{N}_k$ is found by requiring the wave function to be normalized, i.e.

$$1 \stackrel{!}{=} \mathcal{N}_k^2 \sum_{n_2=0}^{\infty} |u_k|^{2n_2}. \tag{9}$$

This geometric series is only convergent if $|u_k|^2 < 1$, which is the case for $\frac{2\pi}{3} < k < \frac{4\pi}{3}$. This means that the edge state wave function is only normalizable for this k interval, and this is consistent with the previous numerical result shown in Fig. 1b, i.e., that the edge state exists only between the Dirac points. In this interval we have $\mathcal{N}_k = \sqrt{1-|u_k|^2}$ and the edge state reads

$$\hat{e}^\dagger_{kB} = \mathcal{N}_k \sum_{n_2=0}^{\infty} (-u_k)^{n_2} \hat{c}^\dagger_{kn_2B}, \quad \frac{2\pi}{3} < k < \frac{4\pi}{3}. \tag{10}$$

Thus, the edge state wave function decays exponentially with n_2, i.e., it is localized at the edge (near $n_2 = 0$). The wave function of the edge states is shown in Fig. 2b. Apparently, for $k = \pi$ the edge states are completely localized at the edge and start to leak into the bulk for $k \neq \pi$. As k approaches the Dirac points at $2\pi/3$ and $4\pi/3$, the edge state delocalizes completely and merges into the bulk bands.

We have discussed a single edge in a half-infinite system and found that there are edge states with exactly zero energy in a certain k-space interval. More interesting (and also more realistic) is the situation with two edges, i.e., a graphene nanoribbon with two opposite zigzag edges. Since we have two equal edges, we expect localized edge states at both of them. In the half-infinite setting we have employed the fact that the energy of the edge states is exactly zero. This is not true in the finite ribbon geometry. Instead, the edge state energies are exponentially small in the ribbon width. This complicates the analytical solution for the edge states considerably. In fact, it is possible but not very enlightening to calculate the edge state wave functions analytically. Thus, we just assume for the moment that we had calculated the edge state wave functions $\phi_{kA}(n_2)$ and $\phi_{kB}(n_2)$ and construct the two edge state operators $\hat{e}^\dagger_{kA}$ and $\hat{e}^\dagger_{kB}$ from those wave functions as before.

We have seen (Fig. 2b) that near $k = \pi$ the edge state wave functions are very well localized so that states at opposite edges are essentially independent. Near $k = 2\pi/3$ or $4\pi/3$ the states extend far into the bulk so that the opposite edge states start to 'overlap'. However, it is not exactly a wave function overlap since the states at opposite edges are supported on different sublattices.[3] Of course the hopping Hamiltonian $\hat{H}_0$, which couples the sublattices then also has finite matrix elements between the states $\hat{e}^\dagger_{kA}$ and $\hat{e}^\dagger_{kB}$.[4] Thus, the edge states $\hat{e}^\dagger_{kA/B}$ cannot be eigenstates of the hopping Hamitonian if the ribbon width is finite. Instead, $\hat{H}_0$ projected onto the single-particle Hilbert space spanned by the edge states $|e_{ks}\rangle = \hat{e}^\dagger_{ks}|\mathrm{vacuum}\rangle$ must be of the form

$$\hat{H}_{0,\mathrm{es}} = {\sum_k}' \sum_{ss'} \hat{e}^\dagger_{ks} \langle e_{ks}|\hat{H}_0|e_{ks'}\rangle \hat{e}_{ks'} = {\sum_k}' t_k e^\dagger_{kA} e_{kB} + H.c., \tag{11}$$

where $t_k = \langle e_{kA}|\hat{H}_0|e_{kB}\rangle$. For each momentum this projected Hamiltonian gives rise to the energies $\epsilon_{k\pm} = \pm|t_k|$, with eigenstates $\hat{e}^\dagger_{kA} \pm \hat{e}^\dagger_{kB}$.[5] The energies and wave functions of those eigenstates, which are simple linear combinations of opposite edge states, can easily be computed by numerically diagonalizing the hopping Hamiltonian $\hat{H}_0$ [Eq. (1)]. From these it is easy to reconstruct t_k and $\hat{e}^\dagger_{ks}$.

[3] This is also true in the finite ribbon geometry.

[4] Note that, because the edge states are supported only on one sublattice, $\hat{H}_0$ cannot give rise to non-zero *diagonal* energies of $\hat{e}^\dagger_{kA/B}$ alone.

[5] We have assumed t_k to be real here, which can easily be accomplished since the original hopping Hamiltonian is real.

Equation (11) is the non-interacting version of an effective theory for the edge states. At this point this theory is in a sense trivial because we have solved the complete problem and obtained the effective theory by cutting out only the edge states from the full solution. But the point is that this can be done easily because we are only concerned with non-interacting Hamiltonians on the way to Eq. (11). The interesting part is the generalization to the theory with electron-electron interactions.

We have seen that a zigzag edge not only leads to the well-known change in boundary conditions, which is usually accounted for by applying a simple zone-folding technique. Instead, it introduces additional states $\hat{e}^\dagger_{ks}$ very close to the Fermi level (at zero energy). The wave functions of these states are localized right at the edge s (remember that we denote the edge by the sublattice at which it terminates). In the light of the introduction, where we have argued that electron-electron interactions are of minor importance because the density of states vanishes, this is alarming. At an edge there are many states right at the charge neutrality point, which of course leads to a strongly enhanced local density of states at the Fermi level. As a consequence, our argument for the irrelevance of the electron-electron interaction breaks down at such a zigzag edge and it is not a good approximation to just neglect it. Thus, we need an interacting theory for the edge states.

3 Interacting edge state theories

3.1 Fermionic edge state theory

In the previous section we have derived a simple non-interacting effective theory for the edge states of a graphene nanoribbon and understood that it is *not* a good approximation to neglect the electron-electron interaction. We thus want to derive the effective interactions of the edge states. For this we start from the Hubbard interaction on the honeycomb lattice

$$\hat{H}_U = U \sum_i \hat{n}_{i\uparrow}\hat{n}_{i\downarrow} = U \sum_i \hat{c}^\dagger_{i\uparrow}\hat{c}_{i\uparrow}\hat{c}^\dagger_{i\downarrow}\hat{c}_{i\downarrow}, \tag{12}$$

for which we want to perform a general basis transformation. The new basis $\hat{d}^\dagger_{kn\sigma} = \sum_i \psi_{kn}(i)\hat{c}^\dagger_{i\sigma}$ will essentially be the eigenbasis of $\hat{H}_0$. In order to streamline the notation we introduce simplified indices $1 = (n_1, k_1)$, $2 = (n_2, k_2)$, ... By direct substitution of the $\hat{c}$ operators expressed in the new basis d one finds for the Hubbard Hamiltonian in the new basis

$$\hat{H}_U = U \sum_{1234} \underbrace{\left[\sum_i \psi^*_1(i)\psi_2(i)\psi^*_3(i)\psi_4(i)\right]}_{=\Gamma_{1234}} \hat{d}^\dagger_{1\uparrow}\hat{d}_{2\uparrow}\hat{d}^\dagger_{3\downarrow}\hat{d}_{4\downarrow}, \tag{13}$$

where we have defined the interaction vertex Γ_{1234} in the general new basis. Note the symmetry $\Gamma_{1234} = \Gamma_{3412}$. In our case it is convenient to partition the eigenbasis of $\hat{H}_0$ into bulk states $\hat{b}^\dagger_{kn\sigma}$ and edge states $\hat{e}^\dagger_{ks\sigma}$.[6] Thus $\hat{H}_U$ consists of many different terms with different numbers of edge- and bulk operators.

Up to this point, all manipulations have been exact. Now we perform the crucial approximation. We assume that in all relevant basis states of the total Fock space, the bulk states are always

[6] The latter are not actually eigenstates of $\hat{H}_0$, but the simple linear combinations $\hat{e}^\dagger_{kA\sigma} \pm \hat{e}^\dagger_{kB\sigma}$ are.

occupied if their energy ϵ_{kn} with respect to $\hat{H}_0$ is negative, and empty if ϵ_{kn} is positive. We will only use matrix elements of $\hat{H}_U$ with respect to this restricted Fock space. This assumption is justified because it essentially assumes that the interaction is irrelevant for the bulk states, a fact for which we have given arguments in the introduction. But one may also check the validity of this assumption with hindsight by directly comparing calculations with and without this assumption [3]. Employing this approximation leads to a dramatic simplification in the interaction Hamiltonian. For instance, if $\hat{d}_{4\downarrow}$ in Eq. (13) is a bulk state, then $\hat{d}^\dagger_{3\downarrow}$ must be *the same* bulk state as well. Otherwise the term has no non-zero matrix elements in the restricted Fock space. As a consequence, the number of terms in the sum in Eq. (13) is greatly reduced. In particular, bulk operators must come in pairs for each spin direction. Furthermore, the term with only bulk operators can be dropped since it is a constant in the restricted Foch space. Thus, we find the truncated Hubbard Hamiltonian

$$\hat{H}_U^{\mathrm{trunc}} = U\left[\sum_{1234}\Gamma^{eeee}_{1234}\hat{e}^\dagger_{1\uparrow}\hat{e}_{2\uparrow}\hat{e}^\dagger_{3\downarrow}\hat{e}_{4\downarrow} + \sum_{123}\Gamma^{eebb}_{1233}\hat{e}^\dagger_{1\uparrow}\hat{e}_{2\uparrow}\hat{b}^\dagger_{3\downarrow}\hat{b}_{3\downarrow} + \sum_{123}\Gamma^{bbee}_{3312}\hat{b}^\dagger_{3\uparrow}\hat{b}_{3\uparrow}\hat{e}^\dagger_{1\downarrow}\hat{e}_{2\downarrow}\right]. \quad (14)$$

The second and third term can be further simplified by employing the above-mentioned symmetry of the vertex Γ. Combining those last two terms we get

$$U\sum_{12,\sigma}\hat{e}^\dagger_{1\sigma}\hat{e}_{2\sigma}\sum_{3}\Gamma^{eebb}_{1233}\hat{b}^\dagger_{3\bar{\sigma}}\hat{b}_{3\bar{\sigma}} = U\sum_{12,\sigma}\hat{e}^\dagger_{1\sigma}\hat{e}_{2\sigma}\sum_{i}\phi^*_1(i)\phi_2(i)\sum_{3}|\psi_3(i)|^2\hat{b}^\dagger_{3\bar{\sigma}}\hat{b}_{3\bar{\sigma}}, \quad (15)$$

where $\bar{\sigma}$ denotes the spin direction opposite to σ. Furthermore, $\phi_*(i)$ are edge state wave functions and $\psi_*(i)$ are bulk state wave functions. In principle, those may be calculated from $\hat{H}_0$, but we will see that it is sufficient to calculate the edge state wave functions only. This is convenient because the number of edge states is much smaller than the number of bulk states. Now consider the sum over the bulk wave functions with the bilinear combination of the bulk state operators. If we apply such a term to the states of the restricted Fock space, the bulk state occupation number $\hat{b}^\dagger_{3\bar{\sigma}}\hat{b}_{3\bar{\sigma}}$ is either 1 or 0, depending on whether ϵ_3 is negative or positive. Thus, as far as the restricted Fock space is concerned, the bulk state sum is nothing but the total density $\rho_b(i)$ of bulk electrons per spin

$$\sum_{3}|\psi_3(i)|^2\hat{b}^\dagger_{3\sigma}\hat{b}_{3\sigma} \overset{\text{rest. Fock sp.}}{\to} \rho_b(i). \quad (16)$$

Furthermore, due to electron-hole symmetry there is a relation between the bulk state density and the edge state density (see Ref. [4])

$$\rho_b(i) = \frac{1}{2} - \frac{1}{2}\sum_{3}|\phi_3(i)|^2 = \frac{1}{2} - \sum_{34}\phi^*_3(i)\phi_4(i)\langle\hat{e}^\dagger_{3\sigma}\hat{e}_{4\sigma}\rangle_0. \quad (17)$$

In the last equality we have used that the average in the non-interacting ground state $\langle\hat{e}^\dagger_{3\sigma}\hat{e}_{4\sigma}\rangle_0 = \delta_{34}\Theta(-\epsilon_3)$, with ϵ_3 a non-interacting edge state energy and that, due to electron-hole symmetry, the total density of the edge state electrons with negative energy equals the one with positive energy. Thus, *we have eliminated even the non-interacting part of the bulk states from the effective edge state theory.* Moreover, we may recombine the wave functions to the interaction

vertex Γ and find a compact form of the truncated Hubbard Hamiltonian

$$\hat{H}_U^{\text{trunc}} = U \sum_{1234} \Gamma^{eeee}_{1234} \left(\hat{e}^\dagger_{1\uparrow} \hat{e}_{2\uparrow} - \langle \hat{e}^\dagger_{1\uparrow} \hat{e}_{2\uparrow} \rangle_0 \right) \left(\hat{e}^\dagger_{3\downarrow} \hat{e}_{4\downarrow} - \langle \hat{e}^\dagger_{3\downarrow} \hat{e}_{4\downarrow} \rangle_0 \right)$$

$$\boxed{= U \sum_{1234} \Gamma_{1234} : \hat{e}^\dagger_{1\uparrow} \hat{e}_{2\uparrow} :: \hat{e}^\dagger_{3\downarrow} \hat{e}_{4\downarrow} := \hat{H}_{U,\text{es}}.} \quad (18)$$

Here we have used the normal order of the electron operators with respect to the non-interacting ground state, which can be conveniently expressed as $: A := A - \langle A \rangle_0$. Furthermore, since the bulk states have vanished from the description, we may also drop the redundant upper indices in the vertex, indicating that only edge state wave functions are involved $\Gamma^{eeee}_{1234} \to \Gamma_{1234}$.
Equation (18) is the effective interacting edge state Hamiltonian for a general edge state basis. It can even be used for edge states at non-zigzag edges, such as chiral edges or even disordered edges. One only must find the proper edge states from the corresponding non-hopping Hamiltonian (see Ref. [3]). Since we restrict ourselves to zigzag edges, however, it is useful to specialize Eq. (18) to the k-space basis of the zigzag edge states $\hat{e}^\dagger_{ks}$ for $s = A, B$. The indices $1, 2, 3, 4$ thus consist of a momentum k and an index s for the edge. But remember that the index $s = A, B$ not only tells us that the edge state is localized at edge A or B; it also tells us that the wave function is non-zero only on one sublattice s. Thus, as we calculate the $\Gamma_{k_1 s_1; k_2 s_2; k_3 s_3; k_4 s_4}$ we only get a non-zero vertex if $s_1 = s_2 = s_3 = s_4$. Furthermore, due to k-conservation, one of the four momentum indices is redundant. So we find

$$\Gamma_{k+q,s_1;k,s_2;k'-q,s_3;k',s_4} = \delta_{s_1 s_2} \delta_{s_1 s_3} \delta_{s_1 s_4} \frac{1}{N_1} \Gamma(k, k', q) \quad (19)$$

$$\Gamma(k, k', q) = \sum_i \phi^*_{k+q,s}(i) \phi_{k,s}(i) \phi^*_{k'-q,s}(i) \phi_{k',s}(i). \quad (20)$$

This means that the effective electron-electron interaction only couples edge states at the same edge. Note also that, due to mirror symmetry at the central axis of a graphene ribbon, $\Gamma(k, k', q)$ is independent of the edge s, so that Eq. (20) can be calculated with any s. The final k-space Hamiltonian then reads

$$\hat{H}_{U,\text{es}} = \frac{U}{N_1} \sum_s {\sum_{k,k',q}}' \Gamma(k, k', q) : \hat{e}^\dagger_{k+qs\uparrow} \hat{e}_{ks\uparrow} :: \hat{e}^\dagger_{k'-qs\downarrow} \hat{e}_{k's\downarrow} :, \quad (21)$$

where the k sum must be restricted such that all momenta are in the interval $[2\pi/3, 4\pi/3]$. Note that this form is very similar to the conventional one-dimensional Hubbard Hamiltonian, the sole difference being the non-unity vertex Γ and the restricted k summation.
Before we continue, let us consider what we have gained up to now by deriving the effective theory. From a computational point of view, we have considerably reduced the number of degrees of freedom. As a result, calculations become much more efficient compared to the full theory. In Section 4.1 this will be demonstrated within a mean-field approximation. But especially for more rigorous calculations that do not rely on the uncontrolled mean-field approximation this effective theory is useful. The fermionic effective theory allows to treat system sizes comparable to or even larger than those of quantum Monte-Carlo calculations by exactly diagonalizing the effective theory. But since exact diagonalization does not suffer from sign problems, the effective theory is more flexible than quantum Monte-Carlo techniques when it comes to adding further terms to the Hamiltonian (see, e.g., Ref [5]). Furthermore, such a theory can be used for completely new approaches which would not be possible within a lattice formulation [7].

3.2 Heisenberg approximation

The fermionic effective theory for the edge states is formulated most conveniently in momentum space. However, if one aims at further reducing the complexity of the theory (at a certain price of accuracy, of course), it is opportune to have a theory in real space at hand. On the basis of such a real-space formulation we will derive an effective theory directly for the spins at the graphene edges. This theory is powerful enough to enable the investigation of extremely large systems with hundreds of millions of lattice sites.

For the real-space formulation we need to perform a certain kind of Fourier transformation. The edge states $\hat{e}^\dagger_{ks\sigma}$ we have used so far can be expressed in terms of the original lattice operators $\hat{c}^\dagger_{n_1n_2s\sigma}$ as

$$\hat{e}^\dagger_{ks\sigma} = \frac{1}{\sqrt{N_1}} \sum_{n_1n_2} e^{ikn_1} \phi_{ks}(n_2) c^\dagger_{n_1n_2s\sigma}, \tag{22}$$

where $\phi_{ks}(n_2)$ is the transverse wave function of the edge state on edge $s = A, B$ and with momentum k along the edge. Remember that we may label the edge by the sublattice it terminates on. It is also important to remember that the momentum for which edge states exist is restricted to one third of the Brillouin zone $2\pi/3 < k < 4\pi/3$. A naive Fourier transformation to real space $\sum_k{}' e^{-ikn_1} \hat{e}^\dagger_{ks\sigma}$ results in a three-fold overcomplete set of edge states. This can most easily be seen by counting the number of k points and the number of n_1 positions in a finite system: there are three times more n_1 points than k points, because the k-space is restricted to one third of its original size. A simple way to resolve this overcompleteness issue is to use the *stretched Fourier transform*

$$\hat{e}^\dagger_{xs\sigma} = \sqrt{\frac{3}{N_1}} \sum_k{}' e^{-3ikx} \hat{e}^\dagger_{ks\sigma}, \tag{23}$$

where $x = 1, 2, \ldots, N_1/3$. It can be seen easily that the edge states $\hat{e}^\dagger_{xs\sigma}$ are orthonormal, i.e., $\{\hat{e}^\dagger_{xs\sigma}, \hat{e}_{x's'\sigma'}\} = \delta_{xx'}\delta_{ss'}\delta_{\sigma\sigma'}$. Furthermore, the states are spatially localized. In fact, they are nothing but Wannier functions obtained from the effective edge state theory in k-space. The wave functions read

$$\phi_{xs}(n_1, n_2) = \frac{\sqrt{3}}{N_1} \sum_k{}' e^{ik(n_1-3x)} \phi_{ks}(n_2) \tag{24}$$

and still have the property that the edge states at edge s live only on the sublattice s at which the edge terminates.

Now we are in a position to transform the effective theory from the previous subsection to the new basis defined by the states $\hat{e}^\dagger_{xs\sigma}$

$$\hat{H}_{\text{es}} = \sum_{xx',\sigma} [t_{xx'} \hat{e}^\dagger_{xA\sigma} \hat{e}_{x'B\sigma} + H.c.] + U \sum_s \sum_{1234} \Gamma_{1234} : \hat{e}^\dagger_{1s\uparrow} \hat{e}_{2s\uparrow} :: \hat{e}^\dagger_{3s\downarrow} \hat{e}_{4s\downarrow} :, \tag{25}$$

where

$$t_{xx'} = \frac{3}{N_1} \sum_k{}' t_k e^{ik(x-x')} \tag{26}$$

$$\Gamma_{1234} = \sum_{n_1n_2} \phi^*_{x_1A}(n_1, n_2) \phi_{x_2A}(n_1, n_2) \phi^*_{x_3A}(n_1, n_2) \phi_{x_4A}(n_1, n_2). \tag{27}$$

Note that the vertex Γ is still s-independent and we have chosen to calculate it from the A-edge wave functions. But of course choosing the B-edge wave functions would not have made any difference.
At this point one might ask what we have gained by transforming to the Wannier basis $\hat{e}^\dagger_{xs\sigma}$. It looks as if the theory had become considerably more complicated. *The central reason for transforming to the Wannier basis is that it provides an obvious route to further approximations.* To understand this, note that in the k-space formulation the vertex $\Gamma(k,k',q)$ is, in an order of magnitude sense, similarly large for all k,k',q. But for $\Gamma_{x_1x_2x_3x_4}$ it is easy to see that if the positions x_i are far apart, the vertex becomes very small compared, say, to the case where all positions are equal. Thus, we now reorganize the sum in the interaction part of Eq. (25).
The terms in which all positions are equal have all the same[7] prefactor $U^* = U\Gamma_{xxxx}$ and read

$$\sum_s U^* \sum_x \hat{n}^e_{xs\uparrow}\hat{n}^e_{xs\downarrow}, \tag{28}$$

where $\hat{n}^e_{xs\sigma} = \hat{e}^\dagger_{xs\sigma}\hat{e}_{xs\sigma} - \frac{1}{2}$. For all realistic geometries and parameter ranges $U^* \approx 0.1U$ is the largest energy scale in the effective theory. And since this term favors the occupation of a Wannier edge state at position x at edge s with exactly one electron, the spin of which will be the most relevant degree of freedom in what follows, it is convenient to formulate the remaining terms in a way that makes the spin physics obvious.
We now consider the terms in the interaction part of Eq. (25) which leave the single occupation of the Wannier states invariant. These are the terms in which the indices in Γ_{1234} are either $1=2$ and $3=4$ or $1=4$ and $2=3$. Further note that $\Gamma_{1122} = \Gamma_{1221}$.[8] With the definition $J^F_{xx'} = U\Gamma_{xxx'x'}$ we may collect all these terms

$$\sum_s \sum_{x\neq x'} J^F_{xx'}\left(\hat{n}^e_{xs\uparrow}\hat{n}^e_{x's\downarrow} + \hat{e}^\dagger_{xs\uparrow}\hat{e}_{x's\uparrow}\hat{e}^\dagger_{x's\downarrow}\hat{e}_{xs\downarrow}\right). \tag{29}$$

Since we are interested in the spin physics (and since it turns out that this is the dominant physics anyway), we define the fermionic representations of the Pauli matrices s^μ with $\mu = x,y,z$

$$\hat{s}^\mu_{xs} = \sum_{\sigma\sigma'} \hat{e}^\dagger_{xs\sigma} s^\mu_{\sigma\sigma'} \hat{e}_{xs\sigma'} \tag{30}$$

and the total electron density $\hat{n}^e_{xs} = \hat{n}^e_{xs\uparrow} + \hat{n}^e_{xs\downarrow} - 1$ in the Wannier orbital x at edge s. With these definitions Eq. (29) takes on the simple form

$$\sum_s \sum_{x<x'} \frac{J^F_{xx'}}{2}\left[\hat{n}^e_{xs}\hat{n}^e_{x's} - \sum_\mu \hat{s}^\mu_{xs}\hat{s}^\mu_{x's}\right] \tag{31}$$

of a ferromagnetic intra-edge coupling $J^F_{xx'}/2$ between the spins of the electrons at sites x and x'. The first density-density term is of minor importance since the effective Hubbard term at the edge, Eq. (28), which is the strongest Hamiltonian in the effective theory, forces each Wannier state to single occupation. Thus, we may safely neglect the density-density term.
What about the many other terms we have neglected? A complete investigation of all those terms is a lengthy but in principle straightforward business. Instead of discussing each subclass

[7] This is due to translational invariance along the edge.
[8] This can be verified directly from the definition in Eq. (27).

of those terms separately, however, let us state the following qualitative argument. We know that the effective Hubbard term at the edges is by far the strongest term and thus we know that there is essentially one electron per Wannier state. All terms that we have not yet taken into account change this single occupation of the Wannier states. And since these terms themselves are much smaller than the effective Hubbard term, we may treat them perturbatively. These terms typically transfer an electron from one site to the other. But due to the Pauli principle this is only possible if there are electrons with opposite spin directions on those two sites. Thus, there will be an energy gain for antiparallel spin alignments, which, as usual, results in an antiferromagnetic coupling in second order perturbation theory between two electron spins in the Wannier states x and x'.[9] However, by the terms in Eq. (31), these two spins are already ferromagnetically coupled – and this ferromagnetic coupling will be much stronger than the perturbative antiferromagnetic coupling coming from the remaining terms. Thus, these terms will only slightly diminish the ferromagnetic intra-edge coupling. We only need to keep this in the back of our mind for possible quantitative considerations later, but for the moment we keep only the effective Hubbard term and the ferromagnetic intra-edge coupling

$$\hat{H}_{\rm es} \approx \sum_{xx',\sigma} [t_{xx'} \hat{e}^\dagger_{xA\sigma} \hat{e}_{x'B\sigma} + H.c.] + \sum_s U^* \sum_x \hat{n}^e_{xs\uparrow} \hat{n}^e_{xs\downarrow} - \sum_s \sum_{x<x'} \frac{J^F_{xx'}}{2} \sum_\mu \hat{s}^\mu_{xs} \hat{s}^\mu_{x's}. \qquad (32)$$

If not for the hopping term (first term) in Eq. (32), we could completely get rid of the fermionic representation and work exclusively with the spin degrees of freedom. So our aim is to reduce also the hopping term to its contribution to the spin physics. For this we consider one pair of Wannier orbitals at different edges (xA) and $(x'B)$ and assume that there are in total two electrons in these two orbitals. Thus, our basis is

$$|\uparrow;\uparrow\rangle,\ \ |\downarrow;\downarrow\rangle,\ \ |\uparrow;\downarrow\rangle,\ \ |\downarrow;\uparrow\rangle,\ \ |\uparrow\downarrow;0\rangle,\ \ |0;\uparrow\downarrow\rangle, \qquad (33)$$

where the last two states are those where one Wannier orbital is doubly occupied. Let us write down the Matrix representation of $\hat{H}_{\rm es}$ in Eq. (32) in this subspace:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -t_{xx'} & -t_{xx'} \\ 0 & 0 & 0 & 0 & t_{xx'} & t_{xx'} \\ 0 & 0 & -t_{xx'} & t_{xx'} & U^* & 0 \\ 0 & 0 & -t_{xx'} & t_{xx'} & 0 & U^* \end{pmatrix} \qquad (34)$$

One can directly check that this matrix has zero-energy eigenstates $|\uparrow;\uparrow\rangle$, $|\downarrow;\downarrow\rangle$, and $(|\uparrow;\downarrow\rangle + |\downarrow;\uparrow\rangle)/\sqrt{2}$, which are exactly the three spin triplets $|T^+\rangle$, $|T^-\rangle$, and $|T^0\rangle$, respectively. The spin singlet $(|\uparrow;\downarrow\rangle - |\downarrow;\uparrow\rangle)/\sqrt{2}$ is not an exact eigenstate of the Hamiltonian – it couples to the doubly occupied states via the hopping terms. But since these doubly occupied states have a much higher energy U^* than all other energy scales, we can treat this coupling perturbatively. However, it is also easily possible to fully diagonalize this matrix analytically. Doing so and expanding the eigenvalues in t_{xx}/U, one finds that the singlet acquires an energy correction

[9] The same effect will give rise to antiferromagnetic inter-edge couplings, which we will discuss in some more detail below. So if you are not familiar with this strong coupling limit of the Hubbard model, just go on reading – it will be explained.

of $-4t^2_{xx'}/U^*$ in leading order. This means that in the low-energy sector, where the Wannier orbitals are singly occupied, the antiparallel alignment of the electron spins (this is the singlet) is favored by an energy $J^A_{xx'} = 4t^2_{xx'}/U^*$.
So we have finally reduced the fermionic edge state theory to the leading spin physics. Now we define our spin-$\frac{1}{2}$ operators S^μ_{xs} for the electron spins in the Wannier states

$$\hat{S}^\mu_{xs} \hat{=} \frac{1}{2}\hat{s}^\mu_{xs}. \tag{35}$$

Note that there is no equal sign in the above equation because $\hat{s}$ acts on the fermionic Fock space, while $\hat{S}$ only acts on a spin space. In terms of these spin operators we may formulate the Heisenberg theory of edge magnetism as

$$\hat{H}_{\mathrm{H}} = \sum_{x,x'} J^A_{xx'} \hat{\mathbf{S}}_{xA} \cdot \hat{\mathbf{S}}_{x'B} - \sum_s \sum_{x<x'} J^F_{xx'} \hat{\mathbf{S}}_{xs} \cdot \hat{\mathbf{S}}_{x's}. \tag{36}$$

Let us again discuss what we have gained by deriving this effective spin theory. Such a Hamiltonian as in Eq. (36) can be simulated very efficiently, e.g., by quantum Monte-Carlo techniques (see Chapter B4) or by the density matrix renormalization group. Computations with tens of thousands of effective spins are possible. Such a number of effective spins corresponds to tens or even hundreds of millions of carbon lattice sites (see, e.g., Ref [6]), which is the size of actual graphene ribbons available in experiments. On the other hand, a comparison of the solution of this effective theory with the exact solution of the lattice system[10] shows that the predictions of the effective theory are extremely close to the results of exact methods, such as quantum Monte-Carlo techniques [3]. Most importantly, however, the effective Heisenberg theory enables the investigation of very fundamental questions regarding the interrelation of classical and quantum physics in magnetic systems. This will be discussed Section 4.3.

4 The nature of edge magnetism

In the previous section we have derived a powerful description of correlation effects at graphene edges. Now we want to actually use those effective theories in order to understand what edge magnetism actually is.

4.1 Mean-field approximation

Mean-field approximations are a simple way of treating interacting systems. One may formulate the aim of mean-field theory as follows: *Find the non-interacting theory which gives the best approximation to an interacting theory.* However, one must emphasize here that mean-field theory is *not* a controlled approximation. Not only that it usually overestimates interaction effects; also it easily develops artifacts and provides qualitatively wrong results. So, mean-field theory must be used with caution! And this is especially true in one dimension as in the case of edge magnetism.
But despite of all its potential issues, mean-field theory is usually a good starting point. Thus, we perform the mean-field approximation of the fermionic edge state theory $\hat{H} = \hat{H}_{0,\mathrm{es}} + \hat{H}_{U,\mathrm{es}}$

[10] This is possible only for small systems, of course.

[see Eqs. (21) and (11)]. The mean-field approximation of $\hat{H}_{U,\mathrm{es}}$ reads[11]

$$\hat{H}_{U,\mathrm{es}} = \frac{U}{N_1}\sum_s {\sum_{k,k',q}}' \Gamma(k,k',q) : \hat{e}^\dagger_{k+qs\uparrow}\hat{e}_{ks\uparrow} :: \hat{e}^\dagger_{k'-qs\downarrow}\hat{e}_{k's\downarrow} : \tag{37}$$

$$\overset{\text{M.F.}}{\approx} \frac{U}{N_1}\sum_s {\sum_{k,k'}}' \Gamma(k,k',0)\left[\hat{e}^\dagger_{ks\uparrow}\hat{e}_{ks\uparrow}\langle: \hat{e}^\dagger_{k's\downarrow}\hat{e}_{k's\downarrow} :\rangle + \langle: \hat{e}^\dagger_{ks\uparrow}\hat{e}_{ks\uparrow} :\rangle\hat{e}^\dagger_{k's\downarrow}\hat{e}_{k's\downarrow}\right] = \hat{H}^{\mathrm{MF}}_{U,\mathrm{es}},$$

where $\langle\cdot\rangle$ is the average with respect to the mean-field Hamiltonian $\hat{H}^{\mathrm{MF}}_{\mathrm{es}} = \hat{H}_{0,\mathrm{es}} + \hat{H}^{\mathrm{MF}}_{U,\mathrm{es}}$. In the second line we have dropped the normal order in the non-averaged fermion bilinears, since this only leads to a constant term in the Hamiltonian. Furthermore, we have assumed that momentum is still a good quantum number in the mean-field theory (which is a very reasonable assumption) so that $\langle: \hat{e}^\dagger_{k+qs\sigma}\hat{e}_{ks\sigma} :\rangle \propto \delta_{q,0}$. Thus, only the $q = 0$ term in the momentum sum survives. Due to the symmetry $\Gamma(k,k',0) = \Gamma(k',k,0)$ the mean-field Hamiltonian may be further simplified to

$$\hat{H}^{\mathrm{MF}}_{U,\mathrm{es}} = U\sum_{s\sigma k}\hat{e}^\dagger_{ks\sigma}\hat{e}_{ks\sigma}\tilde{n}_{s\bar{\sigma}}(k), \tag{38}$$

with

$$\tilde{n}_{s\sigma}(k) = \frac{1}{N_1}\sum_{k'}\Gamma(k,k',0)\langle: \hat{e}^\dagger_{k's\sigma}\hat{e}_{k's\sigma} :\rangle. \tag{39}$$

The total mean-field Hamiltonian $\hat{H}^{\mathrm{MF}}_{\mathrm{es}}$ is bilinear in the edge state operators and diagonal in k. The remaining degrees of freedom are the sublattice s and the spin σ, i.e., $\hat{H}^{\mathrm{MF}}_{\mathrm{es}}$ is a sum of 4×4 blocks. Furthermore, each block consists of two diagonal 2×2 blocks – one for each spin σ. Thus, $\hat{H}^{\mathrm{MF}}_{\mathrm{es}}$ is very easy to diagonalize. One can either perform this diagonalization analytically and obtain a closed formula for the averages $\langle: \hat{e}^\dagger_{k's\sigma}\hat{e}_{k's\sigma} :\rangle$ in terms of the $\tilde{n}_{s\sigma}(k)$ from the previous run, or one can simply diagonalize this 4×4 matrix numerically. Here, we will follow the latter route. We write the Hamiltonian as

$$\hat{H}^{\mathrm{MF}}_{\mathrm{es}} = {\sum_k}'(\hat{e}^\dagger_{kA\uparrow},\hat{e}^\dagger_{kB\uparrow},\hat{e}^\dagger_{kA\downarrow},\hat{e}^\dagger_{kB\downarrow})\underbrace{\begin{pmatrix}\tilde{n}_{A\downarrow}(k) & t_k & 0 & 0\\ t_k & \tilde{n}_{B\downarrow}(k) & 0 & 0\\ 0 & 0 & \tilde{n}_{A\uparrow}(k) & t_k\\ 0 & 0 & t_k & \tilde{n}_{B\uparrow}(k)\end{pmatrix}}_{=\mathbf{h}(k)}\begin{pmatrix}\hat{e}_{kA\uparrow}\\ \hat{e}_{kB\uparrow}\\ \hat{e}_{kA\downarrow}\\ \hat{e}_{kB\downarrow}\end{pmatrix}. \tag{40}$$

The averages $\langle: \hat{e}^\dagger_{k's\sigma}\hat{e}_{k's\sigma} :\rangle$ can be calculated from the eigenvectors $\boldsymbol{\chi}_{kn}$, $n = 1,2,3,4$ of $\mathbf{h}(k)$, which we find numerically. It is assumed that the eigenvectors are ordered such that lower energies correspond to smaller n. We aim at a half-filled system, which means that the two lowest energy eigenstates of $\mathbf{h}(k)$ are occupied for each k.[12] Thus, one finds

$$\langle: \hat{e}^\dagger_{kA\uparrow}\hat{e}_{kA\uparrow} :\rangle = |(\boldsymbol{\chi}_{k1}(k))_1)|^2 + |(\boldsymbol{\chi}_{k2}(k))_1)|^2 - 1/2 \tag{41}$$

$$\langle: \hat{e}^\dagger_{kB\uparrow}\hat{e}_{kB\uparrow} :\rangle = |(\boldsymbol{\chi}_{k1}(k))_2)|^2 + |(\boldsymbol{\chi}_{k2}(k))_2)|^2 - 1/2 \tag{42}$$

$$\langle: \hat{e}^\dagger_{kA\downarrow}\hat{e}_{kA\downarrow} :\rangle = |(\boldsymbol{\chi}_{k1}(k))_3)|^2 + |(\boldsymbol{\chi}_{k2}(k))_3)|^2 - 1/2 \tag{43}$$

$$\langle: \hat{e}^\dagger_{kB\downarrow}\hat{e}_{kB\downarrow} :\rangle = |(\boldsymbol{\chi}_{k1}(k))_4)|^2 + |(\boldsymbol{\chi}_{k2}(k))_4)|^2 - 1/2. \tag{44}$$

[11] We have dropped the constant term in the Hamiltonian.

[12] This is actually a rather nontrivial assumption. However, one may check that without the assumption of 'half filling per k', the results are the same. Only the algorithm becomes a little more complicated.

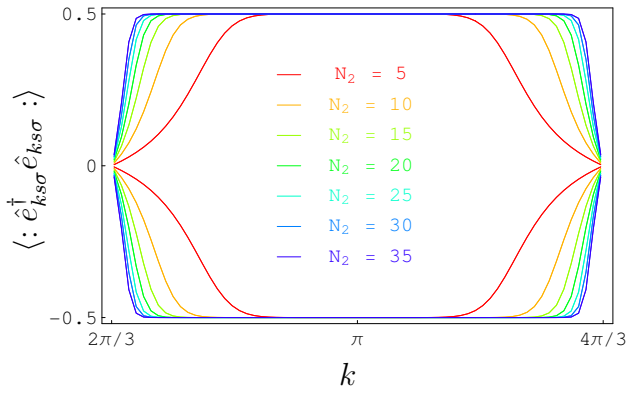


Fig. 3: *Self-consistent mean-field solutions of the averages* $\langle : \hat{e}^\dagger_{ks\sigma}\hat{e}_{ks\sigma} : \rangle$ *for ribbon lengths* $N_1 = 200$ *and different ribbon widths* N_2. *The positive curves correspond to* $(A\uparrow)$, $(B\downarrow)$ *and the negative curves correspond to* $(B\uparrow)$, $(A\downarrow)$. *In the calculations we have assumed* $U = J$.

Given an initial guess for the $\tilde{n}_{s\sigma}(k)$, which enters $\mathbf{h}(k)$, one then calculates the averages $\langle : \hat{e}^\dagger_{k's\sigma}\hat{e}_{k's\sigma} : \rangle$ and from them a new set of $\tilde{n}_{s\sigma}(k)$. This constitutes a self-consistency cycle, which we run until the averages do not change any longer.
Figure 3 shows some solutions for different ribbon widths after the self-consistency cycle has converged. One can see that for all ribbon widths (parametrized by N_2) the different edges A,B have opposite spin polarizations. The limits $\pm 1/2$ correspond to full spin polarization. Further note that, although we have used fixed Hubbard U and hoppings J on the original honeycomb lattice, the opposite spin polarizations at opposite edges are practically independent on these parameters. A closer inspection shows, however, that not for all momenta k the spin polarization is saturated. Instead, close to the two Dirac points at $k = 2\pi/3$ and $k = 4\pi/3$ the spin polarization decreases. This has two reasons: (1) The edge states become more and more delocalized into the bulk close to the Dirac points, and eventually opposite edge states start to overlap and hybridize. But if electrons with different spin directions start to hop (one could also say fluctuate) from one edge to the other, the spin polarization is suppressed. (2) Since the edge states are less localized close to the Dirac points, the effective electron-electron interaction they feel becomes also smaller (remember the argumentation in the Introduction). The interaction, however, is the driving force for the spin polarization, and if the driving force becomes weaker, so does the spin polarization.

4.2 Exact results

In the previous subsection we have discussed the simplest possible approximate solution to the interacting effective edge state theory, i.e., the mean-field solution. The central feature of this solution was a spin polarization, i.e., a magnetization at the edges, with opposite spin directions at opposite edges. This suggests that, if one would measure the local magnetization at an edge, it would be finite. Now we will argue that this is generally not true, but only an artifact of the mean-field approximation. Further ingredients are needed in order to really obtain a static spin polarization at the edges. In this section we employ the Heisenberg theory of edge magnetism (36), which, we have to emphasize this here again, has been derived in a controlled way from the original Hubbard model on the honeycomb lattice $\hat{H}_0 + \hat{H}_U$ [see Eqs. (1) and (12)].
The Heisenberg theory of edge magnetism

$$\hat{H}_{\rm H} = \sum_{x,x'} J^A_{xx'}\hat{\mathbf{S}}_{xA}\cdot\hat{\mathbf{S}}_{x'B} - \sum_s\sum_{x<x'} J^F_{xx'}\hat{\mathbf{S}}_{xs}\cdot\hat{\mathbf{S}}_{x's} \tag{45}$$

has the form of a spin ladder with ferromagnetic leg couplings and antiferromagnetic rung cou-

plings. The only difference to the traditional spin ladders is that not only the nearest-neighbor spins couple, but the coupling is rather long-ranged (see Fig. 4).

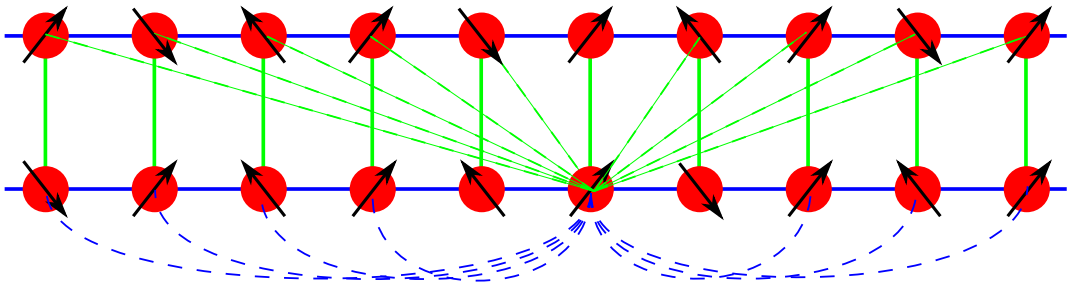


Fig. 4: *Quantum Heisenberg ladder. The green lines represent antiferromagnetic couplings, the blue lines represent ferromagnetic couplings. The solid lines are nearest-neighbor couplings, as they are present in the traditional Heisenberg ladders. The dashed lines represent the additional distant spin couplings present in the Heisenberg theory of edge magnetism. For better visibility, only those distant couplings attached to one particular spin are shown.*

It is well known that the exact ground state of this type of spin ladders is a spin-singlet, i.e., the total spin S of the system, where $S(S+1) = \left(\sum_{s,x} \hat{\mathbf{S}}_{sx}\right)^2$, is zero. As a consequence, the spin polarization $\langle \hat{\mathbf{S}}_{sx} \rangle = 0$ for all s, x,[13] which is in direct contradiction to the mean-field solution. Of course, one may also calculate the spin polarization of the model Hamiltonian (45) directly for small systems by exact diagonalization or by quantum Monte-Carlo techniques and one consistently finds a vanishing spin polarization.
By a quantum Monte-Carlo technique based on the stochastic series expansion, one can simulate (45) for extremely large systems (up to several tens of thousands of spins) and extract the spin-spin correlation $\langle \hat{S}^z_{xs} \hat{S}^z_{x's'} \rangle$ between a given pair of spins. In mean-field theory this correlation is constant – a trivial consequence of the static spin polarization erroneously predicted in this approximation. The exact calculation however shows that the spin-spin correlation decays exponentially as a function of the distance between the two spins under consideration. The length scale on which this decay takes place decreases with the ribbon width (see Fig. 5).
Another useful quantity, which is accessible via the above-mentioned quantum Monte-Carlo technique, is the spin gap Δ. In this context Δ can be thought of as the energy of the lowest excitation above the singlet ground state. Again, mean-field theory predicts $\Delta = 0$ while exact methods give rise to a finite Δ. So in these respects, mean-field theory fails. In particular, it is interesting to note that the failure of mean-field theory is independent of the size of the system, i.e., the length of the nanoribbon.

4.3 Quantum-dynamical aspects

In the previous subsection we have argued that the mean-field prediction of static spin polarizations at the edges of graphene ribbons is wrong, essentially because post mean-field methods, as well as exact theorems [8], consistently predict a unique singlet ground state with zero spin

[13] This can easily be seen by the following hand-waving argument: The ground state is a spin singlet, which means that it is non-degenerate. Furthermore, $\hat{H}_{\mathrm{H}}$ is invariant under global spin rotations. A local spin polarization, however, is not invariant under global spin rotations, which is in contradiction to the fact that a unique ground state must obey the symmetries of the Hamiltonian.

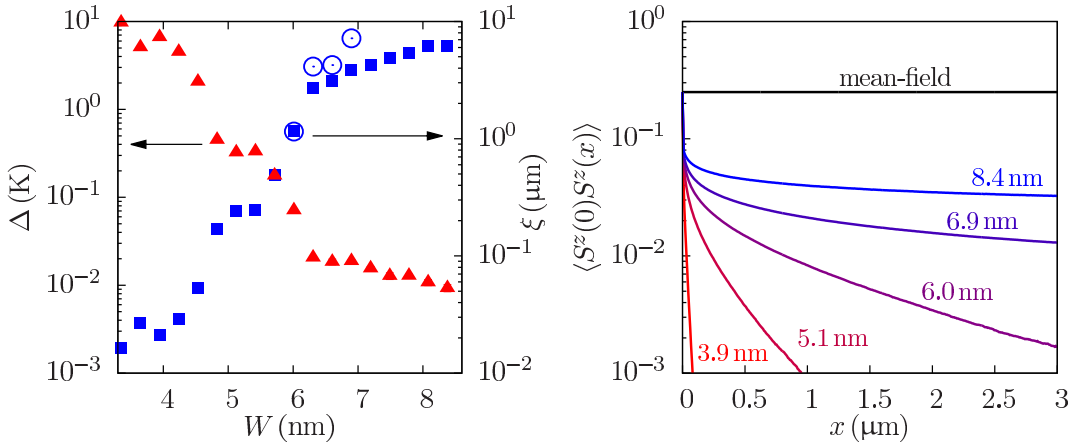


Fig. 5: *Spin gap Δ, correlation length ξ (left) and spin-spin correlation function (right) of the effective Heisenberg ladder corresponding to a chiral ribbon (adopted from Ref. [6]).*

polarization. The problem is that the same argument is in principle applicable to all antiferromagnetic models. Each spin system of finite size with two equally large subgroups of spins which are coupled antiferromagnetically has a spin-singlet ground state and thus, following the above argumentation, can never develop a Néel-like ground state with finite spin polarizations. The good news is that this issue has been known for decades and its resolution for traditional 3D antiferromagnets is well understood by now.

A very convenient model for discussing this issue of the singlet ground state is the Lieb-Mattis model of antiferromagnetism

$$\hat{H}_{\mathrm{LM}} = \frac{2}{N} \sum_{x,x'} \hat{\mathbf{S}}_{xA} \cdot \hat{\mathbf{S}}_{x'B}, \tag{46}$$

where N is the total number of spins per spin group. For simplicity we assume N to be even. This model describes two groups of spins, A and B, and each spin in the A group is coupled antiferromagnetically to each spin in the B group. It is due to this infinite range coupling that this model is exactly solvable. With $\hat{\mathbf{S}}_s = \sum_x \hat{\mathbf{S}}_{xs}$ the Lieb-Mattis Hamiltonian becomes

$$\hat{H}_{\mathrm{LM}} = \frac{2}{N} \hat{\mathbf{S}}_A \cdot \hat{\mathbf{S}}_B = \frac{1}{N} \left[(\hat{\mathbf{S}}_A + \hat{\mathbf{S}}_B)^2 - \hat{\mathbf{S}}_A^2 - \hat{\mathbf{S}}_B^2 \right]. \tag{47}$$

Now we employ the usual spin algebra and see that $\hat{\mathbf{S}}_s^2 = S_s(S_s+1)$ with $S_s = 0, 1, \ldots, N/2$ and $(\hat{\mathbf{S}}_A + \hat{\mathbf{S}}_B)^2 = S(S+1)$ with $S = |S_A - S_B|, \ldots, S_A + S_B$. The lowest possible energy within these degrees of freedom is assumed for minimal S and maximal S_A and S_B, i.e., $S = 0$ and $S_s = N/2$. This is a more formal way of saying that the ground state is a spin singlet with $S = 0$.

Now that we know the exact spectrum $E_{S,S_A,S_B} = [S(S+1) - S_A(S_A+1) - S_B(S_B+1)]/N$ of the Lieb-Mattis model, we may ask under which circumstances the Néel state, where, say, $S_A^z = S_A$ and $S_B^z = -S_B$, is a ground state. The Néel state is a superposition of states with the same $S_{A/B} = N/2$ but with different $S = 0, 1, \ldots$, and is therefore not an eigenstate of $\hat{H}_{\mathrm{LM}}$. But on the other hand the energy differences of the different terms are proportional to $1/N$, i.e., they decrease with the system size N. So in the thermodynamic limit $N \to \infty$ the Néel state indeed becomes an eigenstate. This mechanism is known as the *collapse of the tower of states* [9, 10].

Can we now directly apply this collapse of the tower of states to edge magnetism? This is where the significance of the spin gap Δ becomes apparent, since it directly answers this question with *No!* In the Lieb-Mattis model, as well as in other three-dimensional models of ferromagnetism, the excitation energies of those states needed to build the Néel state approach zero in the thermodynamic limit. But this is not true for the graphene nanoribbons (see Fig. 5). There, the spin gap only depends on the width of the ribbon. Its length is actually irrelevant (as long as the ribbon is longer than the correlation length) for the spin gap.[14] So the relevant excitation energies in nanoribbons do not vanish in the thermodynamic limit [6] and, therefore, the tower of states does not collapse. As a consequence, ribbons of finite width will never have a Néel ground state, even if they are infinitely long.

The spin gap Δ becomes small for very wide ribbons, which raises the question: Shouldn't there be a point at which Δ is so small that it is zero for all practical purposes? Or in other words: *What mechanism renders extremely small excitation energies irrelevant?* In what follows we will tackle this question on a phenomenological level, again employing the exactly solvable Lieb-Mattis model. And at this point it should also be noted that the finite-size Lieb-Mattis model with its infinite-ranged interaction can be simulated by special graphene nanoribbons [6].

We have understood that the classical Néel state $|\psi_N\rangle$ is not an eigenstate, let alone the ground state of the Lieb-Mattis Hamiltonian. This means that a system prepared into $|\psi_N\rangle$ will change with time. Within the Lieb-Mattis model, this time evolution can actually be calculated exactly by straightforward diagonalization of the Hamiltonian (see, e.g., Ref [6]). We will not do this here explicitly, but only state the main result: The sublattice magnetization, which is maximal at the time of the preparation of the system in $|\psi_N\rangle$, decays to zero on a 'quantum-dynamics' time scale $\tau_{\rm qd} \sim \hbar/\Delta$, where Δ is the first excitation energy (the spin gap). Now we assume that at time $t_0 = 0$ we measure the spin polarization of the system (zeroth measurement), with the result that on the sublattice A (or, equivalently, on the edge A) the polarization is, say, positive, while being opposite on the other sublattice. In other words, we have prepared the system in a Néel state via the zeroth measurement. The next measurement takes place at time $t = t_1$. What will we measure? If $t_1 \ll \tau_{\rm qd}$, the system had no time to evolve away from the Néel state. Therefore the result of the first measurement will be the same as the result of the zeroth, with practically 100% probability. If we then proceed with the second, third, ... measurement, always with $t_n - t_{n-1} \ll \tau_{\rm qd}$, the system will stay in the initially prepared Néel state. Our measurement frequency is so large that it *pins* the system to a certain state, no matter if this is an eigenstate of the Hamiltonian or not. Phenomenon is known as the *Quantum Zeno effect* [11]. In the other limit $t_n - t_{n-1} \gg \tau_{\rm qd}$ the quantum dynamics due to the system Hamiltonian is faster than our measurement frequency so that the prepared spin polarization decays completely between successive measurements. In this case successive measurements will not yield the same results but will fluctuate.

This mechanism described above is at the very heart of the more general crossover from quantum physics to classical physics. The most important concept in this field is *decoherence*. Conceptually, decoherence enters the game via the notion that it is not necessarily a dedicated measurement apparatus which observes the system in order to pin it in an actually unstable state (e.g. the Néel state). The environment, which is usually very hard to decouple from the system under consideration, also plays the role of an observer (see, e.g., the relatively new concept of *environment as a witness* in the context of Quantum Darwinism [12]). Thus, usually it is the

[14] This is a property well known for quantum spin ladders.

time scale of decoherence τ_d which must be compared to the time scale of quantum dynamics τ_{qd} in order to determine whether the system is perceived as classical (with spin polarizations and Néel order) or quantum (with a singlet ground state) by an external observer.

5 Conclusion

It was the aim of this chapter to demonstrate how the quantum-mechanical nature of a certain phenomenon, edge magnetism in our case, can be investigated theoretically without resorting to uncontrolled approximations that spoil the true quantum nature of the ground state. The central point of this theory was the reduction of the degrees of freedom to the most relevant ones, thereby enabling the study of realistically large system with hundreds of millions of lattice sites with post-mean-field methods. As a result of this reduction we obtained two effective low-energy theories for the interacting edge states of a graphene nanoribbon.
On the basis of these theories we have developed a systematic understanding of edge magnetism. We have understood that the naive mean-field-theoretical treatment shows severe artifacts which suggest the presence of a static magnetization at zigzag edges of graphene's honeycomb lattice. Exact methods which do not rely on the mean-field approximation, however, show that this simple picture is incomplete in that it does not respect the central role of quantum fluctuations. Taken into account properly, these force the system into a singlet ground state and thereby wash out the magnetization. Nevertheless, by the much simpler effective description we were able to identify concepts and mechanisms (the quantum Zeno effect and decoherence) which potentially destroy the quantum nature of edge magnetism in favor of classical spin polarizations. Is should also be clear from this study that the simple question, if edge magnetism is a classical magnetization effect or rather a quantum effect, cannot be answered from within the isolated theory of a graphene lattice. The role of the environment and of the measurement itself must be accounted for properly.

References

[1] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).

[2] M. Fujita, K. Wakabayashi, K. Nakada, and K. Kusakabe, J. Phys. Soc. Jpn. **65**, 1920 (1996).

[3] M. J. Schmidt, M. Golor, T. C. Lang, and S. Wessel, Phys. Rev. B **87**, 245431 (2013).

[4] M. J. Schmidt and D. Loss, Phys. Rev. B **82**, 085422 (2010).

[5] D. Luitz, F. F. Assaad, and M. J. Schmidt, Phys. Rev. B **83**, 195432 (2011).

[6] M. Golor, S. Wessel, and M. J. Schmidt, arXiv:1309.4446 (2013).

[7] H. Karimi and I. Affleck, Phys. Rev. B **86**, 115446 (2012).

[8] E. H. Lieb, Phys. Rev. Lett. **62**, 1201 (1989).

[9] P. W. Anderson, Phys. Rev. **86**, 694 (1952).

[10] H. Neuberger and T. Ziman, Phys. Rev. B **39**, 2608 (1989).

[11] B. Misra and E. C. G. Sudarshan, Journal of Mathematical Physics **18**, 756 (1977).

[12] W. H. Zurek, Nat. Phys. **5** 181 (2009).

B 7 The functional renormalization group approach to quantum many-body physics[1]

V. Meden
Institut für Theorie der Statistischen Physik
RWTH Aachen University

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercom-

1 Introduction

Many essentially analytical approaches to the quantum many-body problem of model Hamiltonians are based on the idea of a perturbative expansion in a small parameter; in its purest form this is plain perturbation theory [1, 2, 3]. Often the single-particle problem encountered at vanishing two-particle interaction can be solved exactly—although, already this might provide a challenge if e.g. several bands matter, disorder cannot be neglected or effective couplings such as spin-orbit interaction are relevant. Throughout this lecture we will assume that the Green function for $U = 0$ can be computed and that U, being the amplitude of the two-particle interaction, is our small parameter. We aim at the computation of the one-particle irreducible vertex functions such as the free energy (zero-particle vertex), the self-energy (one-particle vertex), and the effective two-particle interaction (two-particle vertex) from which observables, spectral functions, and correlation functions of interest can be obtained. To classify the different terms contributing in the perturbation theory Feynman diagrams become handy. Typically, perturbation theory is limited to the first one or two orders in U as the effort to compute further terms becomes exceedingly large.
Here we do not aim to give a complete account of the limitations of low-order perturbation theory in the two-particle interaction and thus only describe two typical scenarios which motivate the use of nonperturbative approaches.

1) Perturbation theory in U, that is a Taylor expansion of vertex functions around $U = 0$, usually leads to an asymptotic series. Evaluating the n-th order Taylor polynomial beyond some order-dependent characteristic interaction strength $U_{\mathrm{c}}^{(n)}$ typically leads to results which do not even qualitatively agree with the exact ones. E.g. an approximate vertex function becomes negative while the exact one is positive. For a given model $U_{\mathrm{c}}^{(n)}$ is usually not known which renders the use of perturbation theory beyond the rather restricted regime of $U \to 0$ at least "dangerous". In applications one is often interested in results beyond the regime of asymptotically small U. In our discussion of the classical and quantum harmonic oscillator with quartic perturbation we will encounter such a situation. In this case resummations of entire classes of Feynman diagrams of related geometry might be useful. The self-consistent Hartree-Fock (mean-field) approximation, the RPA (random phase approximation), and GW are "classic" approximations schemes based on this idea (see lectures A3 by M. Betziger and A4 by C. Friedrich).

2) Analytical expressions encountered in perturbation theory might diverge if, at temperature $T = 0$, the energy is approaching the Fermi energy ε_{F} and simultaneously the momentum k the Fermi momentum k_{F}, that is if the models infrared (low-energy) limit is probed. This constitutes a severe problem as one is particularly interested in interaction effects at low energy scales at which these are most prominent. Interacting electron systems often exhibit very distinct behavior on different energy scales. Composite objects and collective phenomena emerge at scales far below the bare energy scales of the microscopic Hamiltonian (emergent energy scales). For example, in cuprate high-temperature superconductors one bridges three orders of magnitude from the highest scale, the bare Coulomb interaction, via the intermediate scale of short-range magnetic correlations, down to the lowest scale of d-wave superconductivity and other ordering phenomena. This diversity of scales particularly prominent in low-dimensional systems (e.g. the above mentioned effectively

two-dimensional high-temperature superconductors) is the reason for the divergences appearing in perturbation theory in which all scales are treated at once and within the same approximation. We note in passing that the multitude of scales is also a major obstacle to a straightforward numerical solution of microscopic models, since the most interesting phenomena emerge only at low temperatures and in systems with a large size. It is thus natural to treat degrees of freedom with different energy scales successively, descending step by step from higher to lower scales. This is the main idea behind the renormalization group (RG), which you might have heard about in the context of quantum field theory and/or critical phenomena. Here we will encounter a somewhat weaker form of low-energy (often logarithmic) divergences associated to the appearance of an emergent low-energy scale in our spinless quantum dot model discussed below.

The functional RG (funRG; often also called exact RG) provides a vivid and formally appealing formulation of the RG idea which can directly be applied to microscopic models of (nonrelativistic) many-Fermion systems. It is physically as transparent as Wilson's momentum shell RG but can (at least formally) easily be extended to higher orders. As already indicated funRG can cure the shortcoming 2) of perturbation theory. In addition, it overcomes 1). In fact, "classic" diagrammatic resummation schemes such as RPA can be rederived from funRG. The funRG is flexible and other resummations can be achieved as well. With respect to 1) funRG can be viewed as a "renormalization group enhanced perturbation theory".

RG methods have a long tradition in the theory of interacting Fermi systems. Already in the 1970s, various versions of the RG have been used to deal with infrared singularities arising in one-dimensional Fermi systems [4]. Naturally, RG was also applied to (mostly bosonic) effective field theories describing critical phenomena at continuous classical or quantum phase transitions in interacting Fermi systems [5].

Renormalization group approaches dealing with interacting fermions in arbitrary dimensions d have been developed much later. Due to the extended (not point-like) geometry of the Fermi surface singularity in dimensions $d > 1$, the renormalization group flow cannot be reduced to a finite number of running couplings. However, the main reason for the delayed development of a comprehensive RG approach for interacting Fermi systems in higher dimensions was probably not this difficulty, but rather a lack of motivation. The few infrared singularities appearing in three-dimensional Fermi systems could usually be handled by simple resummations of perturbation theory (e.g. by RPA). Triggered by the issue of non-Fermi liquid behavior in two-dimensional systems, and the related discussion on the validity of perturbation theory, systematic RG approaches to interacting Fermi systems in arbitrary dimensions have been developed by various groups in the early 1990s.

The Wilsonian RG for interacting Fermi systems was popularized by Ref. [6], which contains some of the main ideas in a pedagogical style. A Hamiltonian-based RG interpretation of Fermi liquid theory was presented in Ref. [7], in which not only translation invariant systems but also models for magnetic impurities in metals are discussed. As an alternative to the Wilsonian RG one may also use flow equations for Hamiltonians based on infinitesimal unitary transformations, which make the Hamiltonian successively more diagonal [8]. This approach has been used successfully for quantum impurity models and other systems [9]. A numerical version of the RG concept—the so-called called numerical RG (NRG)—particularly suited for quantum impurity problems was discussed by T. Costi in lecture B3.

We will introduce the funRG considering two toy models: the classical and quantum harmonic oscillator with quartic perturbation. The quartic terms $\propto x^4$ and $\propto \hat{x}^4$, respectively, correspond

to the two-particle interaction in fermionic many-body models (see lecture B2 by S. Jakobs). For the oscillator models plain perturbation theory is of course regular (and you might have encountered it before), that is no divergencies appear. The merits of funRG thus boil down to provide resummations which agree very well to the numerically obtained exact results up to very large interactions—in contrast to plain perturbative results the overall shape of the U-dependence of the vertices is captured favorably. In the past roughly 15 years funRG was applied to different two-, one- and "zero-"dimensional many-body systems. An application of the first type will be discussed by C. Honerkamp (see lecture C5). Here we will focus on its application to "zero"-dimensional problems of quantum dots featuring a few interacting degrees of freedom coupled to noninteracting leads. Studying the single-impurity Anderson model (see lecture B2 by T. Costi) as well as the interacting resonant level model we will discuss the merits and shortcomings of the funRG in describing correlation physics of interacting quantum dots characterized by spin as well as charge fluctuations (see lectures B5 by H. Schoeller and C7 by M. Wegewijs). For both models the funRG approach was extended to nonequilibrium (see lecture B2 by S. Jakobs). We will present results for the bias voltage driven steady state as well as the nonequilibrium relaxation dynamics towards this state for the interacting resonant level model. A comprehensive review of the funRG approach to fermionic many-body systems is given in Ref. [10]

2 The classical anharmonic oscillator

In standard notation the Hamilton function of the harmonic oscillator with quartic perturbation is given by

$$H(x,p) = \frac{p^2}{2m} + \frac{1}{2}m\omega_0 x^2 + Ux^4. \tag{1}$$

Within classical statistical mechanics we might be interested in computing the canonical partition function ($\beta = 1/T$)

$$\frac{\mathcal{Z}_{\rm c}}{\mathcal{Z}_{\rm c,0}} = \frac{1}{\mathcal{Z}_{\rm c,0}} \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx\; e^{-\beta H(x,p)}, \tag{2}$$

where $\mathcal{Z}_{\rm c,0}$ denotes the "noninteracting" partition function with $U = 0$. From $\mathcal{Z}_{\rm c}$ thermodynamic observables can be computed as usual. While the momentum integral is the same in the numerator and denominator and drops out the nongaussian integral over position cannot be given in closed form. This observation already contains the essence of the quantum many-body problem as follows from the functional integral approach to the latter [2] (see also lecture B2 by S. Jakobs). To bring our notation closer to the one used in fermionic systems we rewrite Eq. (2) as

$$\frac{\mathcal{Z}_{\rm c}}{\mathcal{Z}_{\rm c,0}} = \frac{1}{\sqrt{2\pi G^0}} \int_{-\infty}^{\infty} d\,\psi e^{-\frac{\psi^2}{2G^0}} e^{-\frac{g}{4!}\psi^4} \tag{3}$$

with $G^0 = (\beta m\omega_0)^{-1}$ and $\beta U = g/4!$. We used that the position integral of $\mathcal{Z}_{\rm c,0}$ gives $\sqrt{2\pi G^0}$. This partition function can be supplemented to become the generating function of Green functions as well as connected Green functions

$$\mathcal{W}(\theta) = \frac{1}{\sqrt{2\pi G^0}} \int_{-\infty}^{\infty} d\psi\, e^{-\frac{\psi^2}{2G^0}}\, e^{-\frac{g}{4!}\psi^4}\, e^{-\theta\psi}\;, \quad \mathcal{W}^c(\theta) = \ln\left[\mathcal{W}(\theta)\right] \tag{4}$$

with the external source θ which here is merely a real number but in actual many-body problems becomes a complex (bosons) or Grassmann (fermions) field [2]. The present choice of definitions implies that

$$G^0 = \left.\frac{d^2\mathcal{W}_0^{\rm c}}{d\theta^2}\right|_{\theta=0} = \left\langle \psi^2 \right\rangle_0 . \tag{5}$$

As usual the averaging with the noninteracting action is denoted by

$$\langle F(\psi)\rangle_0 = \frac{1}{\sqrt{2\pi G^0}} \int_{-\infty}^{\infty} F(\psi) e^{-\frac{\psi^2}{2G^0}}.$$

with an arbitrary function $F(\psi)$.
Taking the $2m$-th derivative of Eq. (4) with respect to θ and setting $\theta = 0$ leads to the m-particle connected Green functions

$$G_0^{\rm c} = \ln \left\langle e^{-\frac{g}{4!}\psi^4} \right\rangle_0 = \ln I_0, \tag{6}$$

$$G_1^{\rm c} = \left.\frac{d^2\mathcal{W}^{\rm c}}{d\theta^2}\right|_{\theta=0} = \frac{\left\langle \psi^2 e^{-\frac{g}{4!}\psi^4} \right\rangle_0}{\left\langle e^{-\frac{g}{4!}\psi^4} \right\rangle_0} = \frac{I_2}{I_0}, \tag{7}$$

$$G_2^{\rm c} = \left.\frac{d^4\mathcal{W}^{\rm c}}{d\theta^4}\right|_{\theta=0} = \frac{\left\langle \psi^4 e^{-\frac{g}{4!}\psi^4} \right\rangle_0}{\left\langle e^{-\frac{g}{4!}\psi^4} \right\rangle_0} - 3\frac{\left\langle \psi^2 e^{-\frac{g}{4!}\psi^4} \right\rangle_0^2}{\left\langle e^{-\frac{g}{4!}\psi^4} \right\rangle_0^2} = \frac{I_4}{I_0} - 3\frac{I_2^2}{I_0^2}. \tag{8}$$

The integrals I_n are defined by

$$I_n = \frac{1}{\sqrt{2\pi G^0}} \int_{-\infty}^{\infty} d\psi\, \psi^n e^{-\frac{\psi^2}{2G^0}} e^{-\frac{g}{4!}\psi^4}.$$

For a given g they can easily be performed numerically providing us with exact results for the Green functions. Relating to the many-body problem we define a self-energy Σ as

$$G_1^{\rm c} = \frac{1}{[G^0]^{-1} - \Sigma}. \tag{9}$$

We emphasize that what we are eventually interested in is $\mathcal{Z}_{\rm c}/\mathcal{Z}_{\rm c,0}$ or equivalently the interacting part of the "free energy" (up to T) $G_0^{\rm c}$. Thus considering higher order (connected) Green (or vertex functions as introduced next) might on first glance be considered as unnecessary. However, as we will see later these are required to obtain an accurate approximation for the free energy within funRG.
With the definition

$$\phi = -d\mathcal{W}^{\rm c}/d\theta \tag{10}$$

a Legendre transform to the generating function of vertex functions is given by

$$\Gamma(\phi) = -\mathcal{W}^{\rm c}(\theta) - \phi\theta - \frac{1}{2}\left[G^0\right]^{-1}\phi^2,$$

where the one-particle irreducible vertices γ_m are obtained as

$$\gamma_m = d^{2m}\Gamma/d\phi^{2m}\big|_{\phi=0} . \tag{11}$$

This directly implies

$$G_0^{\rm c} = -\gamma_0. \tag{12}$$

With

$$d\Gamma/d\phi = -\theta - \phi \left[G^0\right]^{-1} \tag{13}$$

one obtains

$$\begin{aligned} 1 = \frac{d\phi}{d\phi} &= -\frac{d}{d\phi}\frac{d\mathcal{W}^{\rm c}}{d\theta} = -\frac{d\theta}{d\phi}\frac{d^2\mathcal{W}^{\rm c}}{d\theta^2} \\ &= \left(\frac{d^2\Gamma}{d\phi^2} + \left[G^0\right]^{-1}\right)\frac{d^2\mathcal{W}^{\rm c}}{d\theta^2}. \end{aligned} \tag{14}$$

For $\phi = 0 = \theta$ this leads to

$$G_1^{\rm c} = \frac{1}{\left[G^0\right]^{-1} + \gamma_1},$$

which implies $\gamma_1 = -\Sigma$. Taking the second derivative of Eq. (14) with respect to ϕ and setting the external sources to 0 it follows that

$$G_2^{\rm c} = -\left(G_1^{\rm c}\right)^4 \gamma_2.$$

With these relations γ_0, γ_1, and γ_2 can be computed exactly once the I_n are known.
Before computing the γ_m within funRG we treat the problem at hand in second order perturbation theory and the mean-field approach (the self-consistent Hartree-Fock approximation). As discussed in lecture A3 by M. Betziger the latter is very successfully used to study certain types of many-body problems but fails if correlations matter. In fact, it is tempting to define "correlation effects" as those which are not captured by the mean-field approach. As we will see the mean-field approximation works very well in our toy models but leads to artifacts when studying quantum dots with a few interacting degrees of freedom.

2.1 Perturbation theory and mean-field theory

To determine γ_0, $\gamma_1 = -\Sigma$, and γ_2 up to second order perturbation theory in g we use Wicks theorem, the linked-cluster theorem, and the fact that only one-particle irreducible diagrams contribute to γ_m (see lecture B2 by S. Jakobs). Later we are interested in the case $G^0 = 1$ and therefore set up the perturbation theory only for this choice. We find

$$\gamma_0 = \frac{1}{8}g - \frac{1}{12}g^2 + \mathcal{O}(g^3), \tag{15}$$

$$\gamma_1 = -\Sigma = -\frac{1}{2}g + \frac{5}{12}g^2 + \mathcal{O}(g^3), \tag{16}$$

$$\gamma_2 = g - \frac{3}{2}g^2 + \mathcal{O}(g^3). \tag{17}$$

The g-dependence of the γ_n up to second order are compared to the exact results obtained by numerical integration in Fig. 1. While perturbation theory gives very good results for[2] $g \lessapprox$

[2]The regime in which perturbation theory can be used increases with decreasing order of the vertex function.

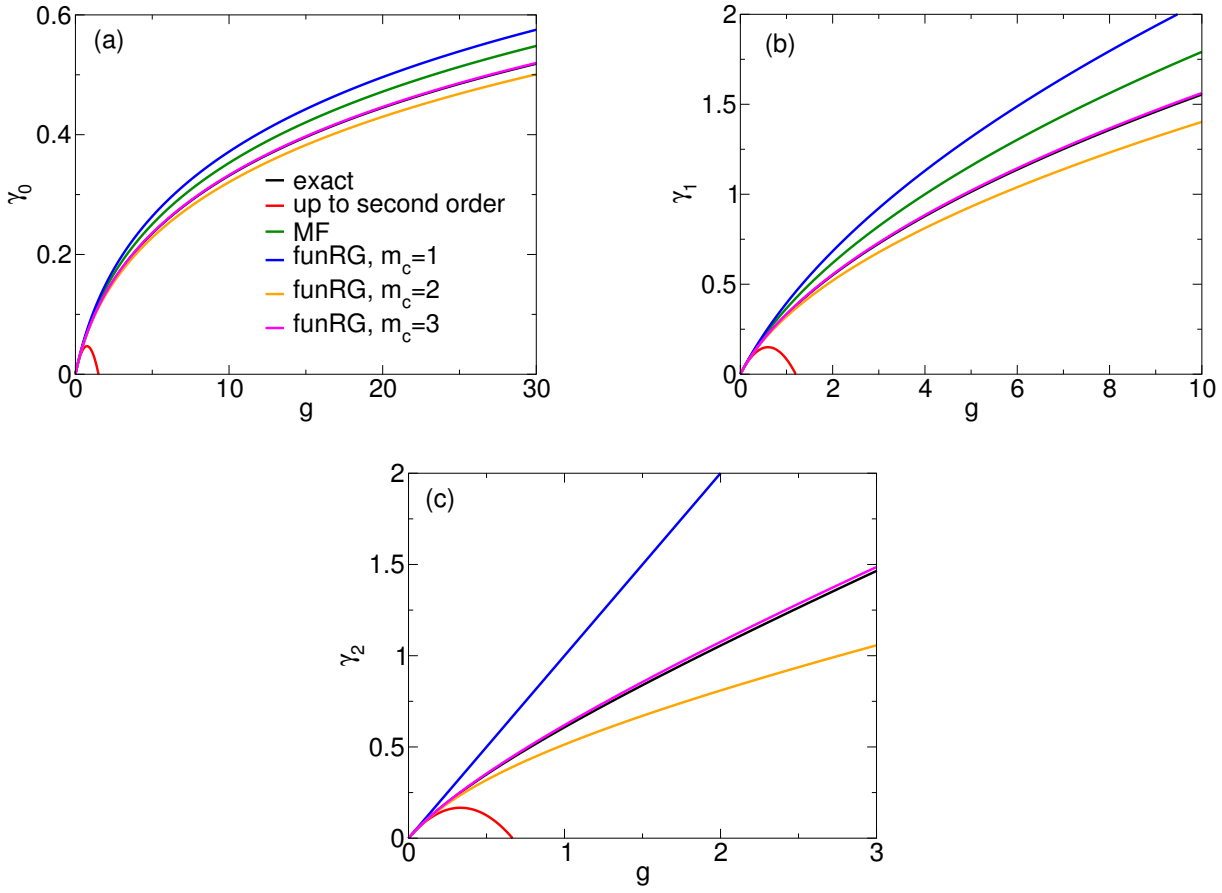


Fig. 1: *The g-dependence of the vertices (a) γ_0, (b) γ_1, and (c) γ_3. The exact results are compared to the ones obtained within second order perturbation theory, the mean-field approximation, and funRG to different truncation orders m_c.*

0.2 the agreement quickly deteriorates beyond this value and perturbation theory fails even qualitatively.

Within the variational mean-field approach one determines the frequency of an effective harmonic oscillator such that it provides a good approximation to the free energy of the oscillator with quartic perturbation. In the quantum case (see below) it is equivalent to the Rayleigh-Ritz variational approach discussed in almost every text book on quantum mechanics. For $G^0 = 1$ the "effective" Hamilton function of the remaining problem reads

$$H(\psi) = H_0(\psi) + H_{\text{int}}(\psi) = \frac{1}{2}\psi^2 + \frac{g}{4!}\psi^4 \tag{18}$$

leading to the mean-field Hamilton function

$$H_{\text{MF}}(\psi) = \frac{1}{2}\psi^2 \left(1 + \frac{g}{2}\left\langle\psi^2\right\rangle_{\text{MF}}\right), \tag{19}$$

with the mean-field expectation value

$$\left\langle\psi^2\right\rangle_{\text{MF}} = \frac{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} d\psi \psi^2 e^{-\psi^2\left(1+g\left\langle\psi^2\right\rangle_{\text{MF}}/2\right)/2}}{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} d\psi e^{-\psi^2\left(1+g\left\langle\psi^2\right\rangle_{\text{MF}}/2\right)/2}} = \frac{1}{1+g\left\langle\psi^2\right\rangle_{\text{MF}}/2}. \tag{20}$$

This equation must be solved self-consistently. In the present toy model the solution of the Hartree-Fock equation can easily be obtained

$$\left\langle \psi^2 \right\rangle_{\mathrm{MF}} = \frac{2}{1+\sqrt{1+2g}} \tag{21}$$

such that the self-energy within the mean-field approximation is given by

$$\Sigma^{\mathrm{MF}} = -\gamma_1^{\mathrm{MF}} = -\frac{1}{2} g \frac{2}{1+\sqrt{1+2g}}. \tag{22}$$

Within this approximation the free energy is as usual (see lecture A3 by M. Betziger) given by the free energy resulting from the mean-field Hamilton function H_{MF} after subtracting the interaction energy

$$\gamma_0^{\mathrm{MF}} = \frac{1}{2} \ln\left[1 + g \left\langle \psi^2 \right\rangle_{\mathrm{MF}} /2\right] - \frac{g}{8} \left\langle \psi^2 \right\rangle_{\mathrm{MF}}^2 . \tag{23}$$

We emphasize that if expanded in g Eqs. (23) and (22) agree to their perturbative counterparts Eqs. (15) and (16) only to linear order in g. We here refrain from discussing the mean-field effective interaction, which would be the two-particle vertex in RPA. In Fig. 1 the g-dependence of $\gamma_{0/1}^{\mathrm{MF}}$ is compared to the exact solution. In contrast to perturbation theory the mean-field solution reproduces the over all shapes of the exact vertex functions although for larger g only qualitatively.

2.2 The funRG flow equations and truncation schemes

We now compute the γ_m within the funRG. In a first step we supplement the noninteracting single-particle Green function G^0 by a parameter Λ leading to $G^{0,\Lambda}$. This must be done in a way such that $G^{0,\Lambda_{\mathrm{i}}} = 0$ and $G^{0,\Lambda_{\mathrm{f}}} = G^0$ ($= 1$ for the chosen parameters; see above) with the initial Λ_{i} and final Λ_{f} values of Λ. Besides this the functional dependence of $G^{0,\Lambda}$ on Λ can be chosen arbitrarily but should be regular. We emphasize that the derivation of the funRG flow equations does not rely on the specific choice. Later we will use the simple dependence $G^{0,\Lambda} = \Lambda$, $\Lambda_{\mathrm{i}} = 0$, $\Lambda_{\mathrm{f}} = 1$ for the classical anharmonic oscillator. For the case of the quantum anharmonic oscillator and fermionic many-body problem of correlated quantum dots we will choose the Λ-dependence of $G^{0,\Lambda}$ in such a way that Λ introduces an infrared cutoff to the problem. This way one can realize the above described successive treatment of energy scales characteristic to a RG procedure. For the classical case the notion of an infrared cutoff does not apply as the variable is energy independent. Via $G^{0,\Lambda}$ the generating functions $\mathcal{W}$, $\mathcal{W}^{\mathrm{c}}$, and Γ become Λ-dependent. By taking the external source $\theta \to \theta^\Lambda$ of $\mathcal{W}^{\mathrm{c},\Lambda}$ as Λ-dependent we can achieve that the (for our purposes) fundamental source $\phi = -d\mathcal{W}^{\mathrm{c},\Lambda}/d\theta^\Lambda$ of Γ^Λ remains Λ-independent. As a crucial step in the RG procedure we now take the derivative

$$\begin{aligned} \frac{d}{d\Lambda}\Gamma^\Lambda(\phi) &= \frac{d}{d\Lambda}\left[-\mathcal{W}^{\mathrm{c},\Lambda}(\theta^\Lambda) - \phi\theta^\Lambda - \frac{1}{2}\left[G^{0,\Lambda}\right]^{-1}\phi^2\right] \\ &= -\dot{\mathcal{W}}^{\mathrm{c},\Lambda}(\theta^\Lambda) - \frac{1}{2}\dot{Q}^\Lambda\phi^2, \end{aligned} \tag{24}$$

were we used Eq. (13) to go from the first to the second line. The dot denotes the derivative with respect to the explicite Λ-dependence and $Q^\Lambda = \left[G^{0,\Lambda}\right]^{-1}$. Taking the derivative of

Eq. (4) (supplemented by all Λ-dependencies) with respect to the explicite Λ-dependence (not the one acquired via θ^Λ) one obtains the funRG flow equation for the generating function of the connected Green functions

$$\dot{\mathcal{W}}^{\mathrm{c},\Lambda} = \frac{1}{2}G^{0,\Lambda}\dot{Q}^\Lambda - \frac{1}{2}\dot{Q}^\Lambda\left[\frac{d^2\mathcal{W}^{\mathrm{c},\Lambda}}{d\theta^2} + \left(\frac{d\mathcal{W}^{\mathrm{c},\Lambda}}{d\theta}\right)^2\right]. \tag{25}$$

Inserting this into Eq. (24) and using Eq. (10) gives

$$\begin{aligned}\dot{\Gamma}^\Lambda &= -\frac{1}{2}G^{0,\Lambda}\dot{Q}^\Lambda + \frac{1}{2}\dot{Q}^\Lambda\left[\frac{d^2\mathcal{W}^{\mathrm{c},\Lambda}}{d\theta^2} + \phi^2\right] - \frac{1}{2}\dot{Q}^\Lambda\phi^2 \\ &= -\frac{1}{2}G^{0,\Lambda}\dot{Q}^\Lambda + \frac{1}{2}\dot{Q}^\Lambda\left[\frac{d^2\Gamma^\Lambda}{d\phi^2} + Q^\Lambda\right]^{-1}.\end{aligned} \tag{26}$$

This partial differential equation constitutes the fundamental flow equation of the funRG approach (in the one-particle irreducible formulation used here). It is supplemented by the initial condition

$$\Gamma^{\Lambda_\mathrm{i}}(\phi) = \frac{g}{4!}\phi^4, \tag{27}$$

which follows form the vanishing of the free propagation at Λ_i, $G^{0,\Lambda_\mathrm{i}} = 0$: the only remaining diagrammatic element is the interaction vertex. We emphasize, that Eq. (26) is still exact, that is, a solution would provide the exact generating function from which the exact vertex functions can be obtained. This implies, that the solution does not depend on the specific choice of the Λ- (cutoff-)dependence of the noninteracting single-particle Green function $G^{0,\Lambda}$ as long as it fulfills the above mentioned boundary conditions at $\Lambda_{\mathrm{i/f}}$.
It will turn out to be efficient to rewrite the last factor of Eq. (26) in terms of the cutoff-dependent full propagator, that is the connected one-particle Green function $G_1^{\mathrm{c},\Lambda} = [Q^\Lambda + \gamma_1^\Lambda]^{-1}$

$$\left[\frac{d^2\Gamma^\Lambda}{d\phi^2} + Q^\Lambda\right]^{-1} = \left[\left(G_1^{\mathrm{c},\Lambda}\right)^{-1} + \frac{d^2\Gamma^\Lambda}{d\phi^2} - \gamma_1^\Lambda\right]^{-1} = \left[1 + G_1^{\mathrm{c},\Lambda}\mathcal{U}\right]^{-1} G_1^{\mathrm{c},\Lambda} \tag{28}$$

with

$$\mathcal{U} = \frac{d^2\Gamma^\Lambda}{d\phi^2} - \gamma_1^\Lambda = \mathcal{O}(\phi^2). \tag{29}$$

Inserting Eq. (28) into Eq. (26) gives

$$\begin{aligned}\dot{\Gamma}^\Lambda(\phi) &= -\frac{1}{2}G^{0,\Lambda}\dot{Q}^\Lambda + \frac{1}{2}G_1^{\mathrm{c},\Lambda}\dot{Q}^\Lambda\left(1 + G_1^{\mathrm{c},\Lambda}\mathcal{U}\right)^{-1} \\ &= \frac{1}{2}\left[G_1^{\mathrm{c},\Lambda} - G^{0,\Lambda}\right]\dot{Q}^\Lambda - \frac{1}{2}G_1^{\mathrm{c},\Lambda}\dot{Q}^\Lambda G_1^{\mathrm{c},\Lambda}\mathcal{U} + \frac{1}{2}G_1^{\mathrm{c},\Lambda}\dot{Q}^\Lambda G_1^{\mathrm{c},\Lambda}\mathcal{U}G_1^{\mathrm{c},\Lambda}\mathcal{U} + \ldots \\ &= \frac{1}{2}\left[G_1^{\mathrm{c},\Lambda} - G^{0,\Lambda}\right]\dot{Q}^\Lambda - \frac{1}{2}\mathcal{S}^\Lambda\mathcal{U} + \frac{1}{2}\mathcal{S}^\Lambda\mathcal{U}G_1^{\mathrm{c},\Lambda}\mathcal{U} + \ldots.\end{aligned} \tag{30}$$

Besides the scale dependent full propagator $G_1^{\mathrm{c},\Lambda}$ the (so-called single-scale) propagator $\mathcal{S}^\Lambda = G_1^{\mathrm{c},\Lambda}\dot{Q}^\Lambda G_1^{\mathrm{c},\Lambda}$ appears.

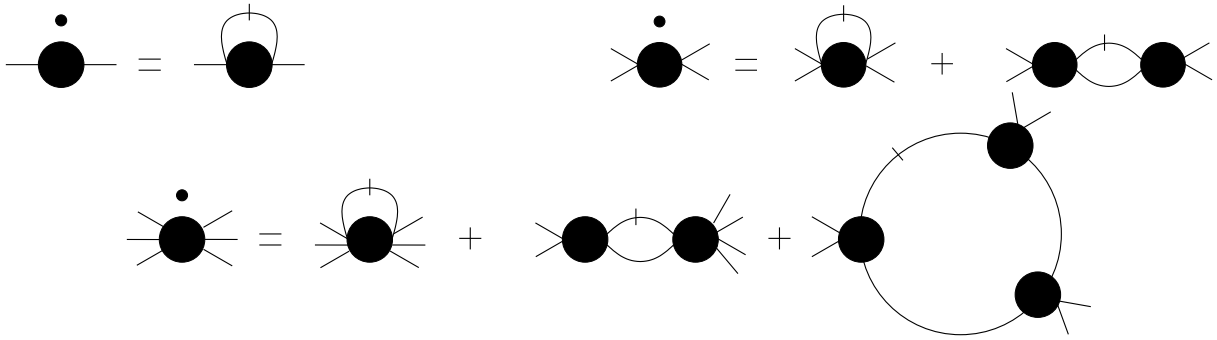


Fig. 2: *Graphical representation of the funRG flow equations for γ_1, γ_2, and γ_3. The line stands for propagation with $G_1^{\mathrm{c},\Lambda}$, the slashed one for propagation with $\mathcal{S}^\Lambda$.*

Instead of attempting to directly solve the flow equation for Γ^Λ in one of the above given forms[3] we will expand Γ^Λ as a Taylor polynomial in the source ϕ

$$\Gamma^\Lambda(\phi) = \sum_{m=0}^{\infty} \frac{1}{(2m)!} \gamma_m^\Lambda \phi^{2m}, \tag{31}$$

with the expansion coefficients being the cutoff-dependent vertices γ_m^Λ. By definition Γ^Λ is an even function of ϕ. Thus only even powers contribute. By comparing the orders of ϕ on the left and right hand side of Eq. (30) one obtains an infinite hierarchy of coupled differential equations for the vertices. The first few read

$$\dot{\gamma}_0^\Lambda = \frac{1}{2}\left[G_1^{\mathrm{c},\Lambda} - G^{0,\Lambda}\right]\dot{Q}^\Lambda, \quad \gamma_0^{\Lambda_\mathrm{i}} = 0, \tag{32}$$

$$\dot{\gamma}_1^\Lambda = -\frac{1}{2}\mathcal{S}^\Lambda \gamma_2^\Lambda, \quad \gamma_1^{\Lambda_\mathrm{i}} = 0, \tag{33}$$

$$\dot{\gamma}_2^\Lambda = -\frac{1}{2}\mathcal{S}^\Lambda \gamma_3^\Lambda + 3\mathcal{S}^\Lambda \gamma_2^\Lambda G_1^{\mathrm{c},\Lambda} \gamma_2^\Lambda, \quad \gamma_2^{\Lambda_\mathrm{i}} = g, \tag{34}$$

$$\dot{\gamma}_3^\Lambda = -\frac{1}{2}\mathcal{S}^\Lambda \gamma_4^\Lambda + 15\mathcal{S}^\Lambda \gamma_2^\Lambda G_1^{\mathrm{c},\Lambda} \gamma_3^\Lambda - 45\mathcal{S}^\Lambda \gamma_2^\Lambda G_1^{\mathrm{c},\Lambda} \gamma_2^\Lambda G_1^{\mathrm{c},\Lambda} \gamma_2^\Lambda, \quad \gamma_3^{\Lambda_\mathrm{i}} = 0. \tag{35}$$

A graphical presentation of these equations reminiscent of Feynman diagrams is given in Fig. 2. For $m \geq 1$ the flow of γ_m^Λ only depends on the γ_n^Λ with $n \leq m$ and on γ_{m+1}^Λ. Remind that $\mathcal{S}^\Lambda$ and $G_1^{\mathrm{c},\Lambda}$ contain the flowing self-energy $-\gamma_1^\Lambda$. The zero-particle vertex γ_0^Λ does not enter on the right hand side of any of the flow equations while its flow is determined by γ_1^Λ. The latter implies that, as mentioned above, to obtain a funRG result for $\gamma_0^{\Lambda_\mathrm{f}}$ we need to consider higher order vertices as well.

To obtain a practical scheme the hierarchy has to be truncated such that the system of equations becomes finite and closes. The guiding principle is perturbation theory. Using that for $m \geq 3$, γ_m has contributions only to order g^m and higher, one sees that the truncation scheme to order m_c in which $\gamma_{m_\mathrm{c}+1}^\Lambda \to \gamma_{m_\mathrm{c}+1}^{\Lambda_\mathrm{i}}$ reproduces all the vertices with $m \leq m_\mathrm{c}$ correctly at least to order m_c (in perturbation theory in g). The flow equations for the γ_m with $m > m_\mathrm{c}$ do not have to be considered. Further down I will show this explicitely. Before doing so we now return to our specific choice of the cutoff function. Inserting that $G^{0,\Lambda} = \Lambda$ and using the definition of $\mathcal{S}^\Lambda$

[3] This might be a reasonable approach in the present toy problem but clearly becomes useless in case of quantum many-body problems for which the corresponding flow equation is a partial functional differential equation.

the flow equations (32)-(35) to truncation order $m_c = 3$ read

$$\dot{\gamma}_0^\Lambda = \frac{1}{2}\frac{\gamma_1^\Lambda}{1+\Lambda\gamma_1^\Lambda}, \tag{36}$$

$$\dot{\gamma}_1^\Lambda = \frac{1}{2}\frac{\gamma_2^\Lambda}{(1+\Lambda\gamma_1^\Lambda)^2}, \tag{37}$$

$$\dot{\gamma}_2^\Lambda = \frac{1}{2}\frac{\gamma_3^\Lambda}{(1+\Lambda\gamma_1^\Lambda)^2} - 3\Lambda\frac{\left(\gamma_2^\Lambda\right)^2}{(1+\Lambda\gamma_1^\Lambda)^3}, \tag{38}$$

$$\dot{\gamma}_3^\Lambda = -15\Lambda\frac{\gamma_2^\Lambda\gamma_3^\Lambda}{(1+\Lambda\gamma_1^\Lambda)^3} + 45\Lambda^2\frac{\left(\gamma_2^\Lambda\right)^3}{(1+\Lambda\gamma_1^\Lambda)^4}, \tag{39}$$

with the initial conditions

$$\gamma_2^{\Lambda_i=0} = g\ , \quad \gamma_m^{\Lambda_i=0} = 0 \ \text{for}\, m \neq 2. \tag{40}$$

They can easily be solved on a computer using standard algorithms. In Fig. 1 the results for the g-dependence of γ_0, γ_1, and γ_2 at the end of the flow at $\Lambda_f = 0$ are compared to the exact ones. The three truncation orders $m_c = 1, 2, 3$ are considered. Similar to the mean-field approximation the over all interaction dependence of the vertices is reproduced quite well. With increasing truncation order the quality of the approximation systematically improves. Already for $m_c = 2$ funRG is superior to the mean-field approximation and for $m_c = 3$ the results are highly accurate at least for the g-ranges shown. The lower the order of the vertex function the longer the approximation is quantitative; note the different g-axis scales. This exemplifies that the feedback of the flowing vertices on the right hand side of the differential RG equations leads to a systematic and meaningful resummation of Feynman diagrams; in this sense funRG is a "renormalization group enhanced perturbation theory". Note that none of the above funRG truncation schemes is equivalent to the mean-field approximation. This will become important further down when discussing quantum dots. In such systems self-consistent Hartee-Fock leads to artificial spontaneous symmetry breaking which can be avoided using funRG.

We emphasize that after truncation the results at fixed m_c will generically depend on the specific choice of the cutoff; while agreeing to order m_c two different cutoff choices lead to different contributions of higher order terms. To gain insights into this dependence one can study the same problem using two or more different cutoff schemes and compare results.

To complete the considerations of the classical harmonic oscillator with quartic perturbation we show how the perturbative results Eq. (16) and (17) follow from the RG flow equations. One expands the γ_m^Λ in powers of g

$$\gamma_m^\Lambda = \sum_{n=1}^{\infty} \gamma_{m,n}^\Lambda g^n,$$

substitutes this expansion on both sides of the flow equations, and compares the orders of g. To first order one obtains

$$\dot{\gamma}_{1,1}^\Lambda = \frac{1}{2}\gamma_{2,1}^\Lambda,$$
$$\dot{\gamma}_{m,1}^\Lambda = 0 \quad \text{for}\ m \geq 2$$

and to second order

$$\dot{\gamma}^{\Lambda}_{1,2} = \frac{1}{2}\gamma^{\Lambda}_{2,2} - \Lambda\gamma^{\Lambda}_{2,1}\gamma^{\Lambda}_{1,1}, \tag{41}$$

$$\dot{\gamma}^{\Lambda}_{2,2} = -3\Lambda\left(\gamma^{\Lambda}_{2,1}\right)^2, \tag{42}$$

$$\dot{\gamma}^{\Lambda}_{m,2} = 0 \quad \text{for } m \geq 3. \tag{43}$$

As the γ^{Λ}_m are one-particle irreducible all vertices with $m \geq 3$ are of higher order in g, or more precisely

$$\gamma^{\Lambda}_{m,n} = 0 \;\text{ for }\; n < m.$$

They do not contribute to the flow equations up to order g^2.
Starting from $\gamma^{\Lambda_{\rm i}=0}_{2,1} = 1$ and $\gamma^{\Lambda_{\rm i}=0}_{n,m} = 0$ the differential equations resulting from the expansion in orders of g can successively be integrated analytically. One first integrates $\dot{\gamma}^{\Lambda}_{2,1}$ and substitutes the result in the equation for $\dot{\gamma}^{\Lambda}_{1,1}$. This way one gets

$$\begin{aligned}\gamma^{\Lambda}_2 &= g + \mathcal{O}(g^2)\\ \gamma^{\Lambda}_1 &= \frac{1}{2}\Lambda g + \mathcal{O}(g^2)\,,\end{aligned}$$

that is for $\Lambda = 1$ we reproduce the results of first order perturbation theory Eqs. (16) and (17). After this step the right hand side of the equation for $\dot{\gamma}^{\Lambda}_{2,2}$ is known and the differential equation can be integrated. This leads to the right hand side of the equation for $\dot{\gamma}^{\Lambda}_{1,2}$. This way one can compute γ_1 and γ_2 up to second order in g

$$\gamma_2 = g - \frac{3}{2}g^2 + \mathcal{O}(g^3) \tag{44}$$

$$\gamma_1 = \frac{1}{2}g - \frac{5}{12}g^2 + \mathcal{O}(g^3)\,. \tag{45}$$

This explicitely shows that one can reproduce second order perturbation theory from truncated funRG (this includes γ_0 which was not discussed here). As elaborated on above using funRG to truncation order $m_{\rm c}$ one systematically captures all terms of plain perturbation theory to order $m_{\rm c}$. Due to the feedback of the vertices on the right hand sides and the differential structure of the RG equations specific higher order terms are resummed. Apparently, the selection inherent to the funRG formalism is meaningful (at least for the present example).

3 A bit of the general formalism

Before studying our second toy problem, namely the quantum harmonic oscillator with quartic perturbation, it is advantageous to take a brief look at the general funRG formalism for quantum many-body problems. Compared to the above the variables and external sources become functions, the generating functions become functionals, and Green and vertex functions become matrices.
Expressed as a (coherent state) functional integral the grand canonical partition function of a system of quantum mechanical particles (either fermions or bosons) interacting via a two-particle potential can be written as [2] (see also lecture B2 by S. Jakobs)

$$\frac{\mathcal{Z}}{\mathcal{Z}_0} = \frac{1}{\mathcal{Z}_0}\int \mathcal{D}\bar{\psi}\psi \exp\left\{(\bar{\psi}, \left[G^0\right]^{-1}\psi) - A_{\rm int}\left(\{\bar{\psi}\},\{\psi\}\right)\right\}, \tag{46}$$

with the interacting part of the action

$$A_{\rm int}\left(\{\bar\psi\},\{\psi\}\right)=\frac{1}{4}\sum_{k_1',k_2',k_1,k_2}\bar v_{k_1',k_2',k_1,k_2}\bar\psi_{k_1'}\bar\psi_{k_2'}\psi_{k_2}\psi_{k_1}, \tag{47}$$

and $\mathcal{Z}_0$ being the noninteracting partition function. Here ψ and $\bar\psi$ denote either Grassmann (fermions) or complex (bosons) fields. The multi-indices $k_j^{(\prime)}$ stand for the quantum numbers of the single-particle basis in which the problem is considered (e.g. momenta and spin directions) and Matsubara frequencies ω. We have introduced the short hand notation

$$\left(\bar\psi,\left[G^0\right]^{-1}\psi\right)=\sum_{k,k'}\bar\psi_k\left[G^0\right]^{-1}_{k,k'}\psi_{k'}\ ,$$

with the propagator G^0 of the related noninteracting problem given as a matrix. The anti-symmetrized (fermions) or symmetrized (bosons) matrix elements of the two-particle interaction are denoted by $\bar v_{k_1',k_2',k_1,k_2}$. They contain the energy conserving factor $\delta_{\omega+\omega',\nu+\nu'}$ and $1/\beta$. The generating functional of the m-particle Green function is given by

$$\mathcal{W}\left(\{\bar\theta\},\{\theta\}\right)=\frac{1}{\mathcal{Z}_0}\int\mathcal{D}\bar\psi\psi\exp\left\{\left(\bar\psi,\left[G^0\right]^{-1}\psi\right)-A_{\rm int}(\{\bar\psi\},\{\psi\})-\left(\bar\psi,\theta\right)-\left(\bar\theta,\psi\right)\right\}, \tag{48}$$

with $\left(\bar\psi,\theta\right)=\sum_k\bar\psi_k\theta_k$ and the external source fields θ and $\bar\theta$. From this the generating functional of the connected m-particle Green function follows as

$$\mathcal{W}^{\rm c}\left(\{\bar\theta\},\{\theta\}\right)=\ln\left[\mathcal{W}\left(\{\bar\theta\},\{\theta\}\right)\right]. \tag{49}$$

The (connected) m-particle Green function $G_m^{(\rm c)}$ can be obtained by taking functional derivatives

$$G_m^{(\rm c)}\left(k_1',\ldots,k_m';k_1,\ldots,k_m\right)=\frac{\delta^m}{\delta\bar\theta_{k_1'}\ldots\delta\bar\theta_{k_m'}}\frac{\delta^m}{\delta\theta_{k_m}\ldots\delta\theta_{k_1}}\mathcal{W}^{(\rm c)}\left(\{\bar\theta\},\{\theta\}\right)\Big|_{\theta=0=\bar\theta}. \tag{50}$$

By a Legendre transformation ($\phi_k=-\delta\mathcal{W}^{\rm c}\left(\{\bar\theta\},\{\theta\}\right)/\delta\bar\theta_k$, $\bar\phi_k=-\zeta\delta\mathcal{W}^{\rm c}\left(\{\bar\theta\},\{\theta\}\right)/\delta\theta_k$)

$$\Gamma\left(\{\bar\phi\},\{\phi\}\right)=-\mathcal{W}^{\rm c}\left(\{\bar\theta\},\{\theta\}\right)-\left(\bar\phi,\theta\right)-\left(\bar\theta,\phi\right)+\left(\bar\phi,\left[G^0\right]^{-1}\phi\right)\ , \tag{51}$$

the generating functional of the one-particle irreducible vertex functions γ_m, with external source fields ϕ and $\bar\phi$ and

$$\gamma_m\left(k_1',\ldots,k_m';k_1,\ldots,k_m\right)=\frac{\delta^m}{\delta\bar\phi_{k_1'}\ldots\delta\bar\phi_{k_m'}}\frac{\delta^m}{\delta\phi_{k_m}\ldots\delta\phi_{k_1}}\Gamma\left(\{\bar\phi\},\{\phi\}\right)\Big|_{\phi=0=\bar\phi}, \tag{52}$$

is obtained. The relations between the $G_m^{(\rm c)}$ and γ_m can be found in text books [2]. They are matrix generalization of the ones discussed in the last section. The zero-particle vertex γ_0 provides the interacting part of the grand canonical potential Ω

$$\Omega=-T\ln\mathcal{Z}=T\gamma_0-T\ln\mathcal{Z}_0\ .$$

For the one-particle Green function we obtain

$$G_1(k';k)=G_1^{\rm c}(k';k)=-\zeta G_{k',k}=\left[\gamma_1-\zeta\left[G^0\right]^{-1}\right]^{-1}_{k',k}\ ,$$

where

$$G_{k',k} = \left[\left[G^0 \right]^{-1} - \Sigma \right]^{-1}_{k',k} ,$$

with the self-energy Σ, and $\zeta = -1$ for fermions or $\zeta = 1$ for bosons, respectively. This implies the relation $\Sigma = \zeta \gamma_1$.
Following the steps described in the last section we introduce a cutoff Λ to the noninteracting propagator $G^0 \to G^{0,\Lambda}$ obeying the boundary conditions given above at $\Lambda_{\mathrm{i/f}}$. The generating functionals, Green functions, and vertices acquire a Λ-dependence. In most of our applications we use a sharp Matsubara frequency cutoff with

$$G^{0,\Lambda} = \Theta \left(|\omega| - \Lambda \right) G^0 \tag{53}$$

and consider $\Lambda_{\mathrm{i}} \to \infty$ as well as $\Lambda_{\mathrm{f}} = 0$. By excluding the degrees of freedom close to the Fermi surface (at $\omega = 0$) problems which are infrared divergent in perturbation theory are regularized; energy scales are treated successively. We differentiate $\mathcal{W}^{\mathrm{c},\Lambda}$ with respect to Λ, which after straightforward algebra leads to

$$\frac{d}{d\Lambda} \mathcal{W}^{\mathrm{c},\Lambda} = \zeta \ \mathrm{Tr} \left(\mathcal{Q}^\Lambda G^{0,\Lambda} \right) + \mathrm{Tr} \left(\mathcal{Q}^\Lambda \frac{\delta^2 \mathcal{W}^{\mathrm{c},\Lambda}}{\delta \bar{\theta} \delta \theta} \right) + \zeta \left(\frac{\delta \mathcal{W}^{\mathrm{c},\Lambda}}{\delta \theta}, \mathcal{Q}^\Lambda \frac{\delta \mathcal{W}^{\mathrm{c},\Lambda}}{\delta \bar{\theta}} \right) , \tag{54}$$

with

$$\mathcal{Q}^\Lambda = \frac{d}{d\Lambda} \left[G^{0,\Lambda} \right]^{-1} . \tag{55}$$

From Eq. (51) we obtain

$$\frac{d}{d\Lambda} \Gamma^\Lambda \left(\{\bar{\phi}\}, \{\phi\} \right) = - \frac{d}{d\Lambda} \mathcal{W}^{\mathrm{c},\Lambda} \left(\{\bar{\theta}^\Lambda\}, \{\theta^\Lambda\} \right) - \left(\bar{\phi}, \frac{d}{d\Lambda} \theta^\Lambda \right) - \left(\frac{d}{d\Lambda} \bar{\theta}^\Lambda, \phi \right) + \left(\bar{\phi}, \mathcal{Q}^\Lambda \phi \right) .$$

Applying the chain rule and using Eq. (54) this leads to

$$\frac{d}{d\Lambda} \Gamma^\Lambda = - \zeta \ \mathrm{Tr} \left(\mathcal{Q}^\Lambda G^{0,\Lambda} \right) - \mathrm{Tr} \left(\mathcal{Q}^\Lambda \frac{\delta^2 \mathcal{W}^{\mathrm{c},\Lambda}}{\delta \bar{\theta}^\Lambda \delta \theta^\Lambda} \right) .$$

Using the standard relation [2] between the second functional derivatives of Γ and $\mathcal{W}^{\mathrm{c}}$ we obtain the functional partial differential equation

$$\frac{d}{d\Lambda} \Gamma^\Lambda = - \zeta \ \mathrm{Tr} \left(\mathcal{Q}^\Lambda G^{0,\Lambda} \right) - \mathrm{Tr} \left(\mathcal{Q}^\Lambda \mathcal{V}^{1,1}_{\bar{\phi},\phi} (\Gamma^\Lambda, G^{0,\Lambda}) \right) , \tag{56}$$

where $\mathcal{V}^{1,1}_{\bar{\phi},\phi}$ stand for the upper left block of the matrix

$$\mathcal{V}_{\bar{\phi},\phi} (\Gamma^\Lambda, G^\Lambda) = \begin{pmatrix} \frac{\delta^2 \Gamma^\Lambda}{\delta \bar{\phi} \delta \phi} - \zeta \left[G^{0,\Lambda} \right]^{-1} & \frac{\delta^2 \Gamma^\Lambda}{\delta \bar{\phi} \delta \bar{\phi}} \\ \frac{\delta^2 \Gamma^\Lambda}{\delta \phi \delta \phi} & \frac{\delta^2 \Gamma^\Lambda}{\delta \phi \delta \bar{\phi}} - \left[\left[G^{0,\Lambda} \right]^{-1} \right]^t \end{pmatrix}^{-1} \tag{57}$$

and the upper index t denotes the transposed matrix.
To obtain differential equations for the γ^Λ_m which include self-energy corrections we express $\mathcal{V}_{\bar{\phi},\phi}$ in terms of G^Λ instead of $G^{\Lambda,0}$. This is achieved by defining

$$\mathcal{U}_{\bar{\phi},\phi} = \frac{\delta^2 \Gamma^\Lambda}{\delta \bar{\phi} \delta \phi} - \gamma^\Lambda_1$$

and using

$$G^{\Lambda} = \left[\left[G^{0,\Lambda} \right]^{-1} - \zeta \gamma_1^{\Lambda} \right]^{-1} \tag{58}$$

which leads to

$$\frac{d}{d\Lambda} \Gamma^{\Lambda} = -\zeta \ \mathrm{Tr} \ \left(\mathcal{Q}^{\Lambda} G^{0,\Lambda} \right) + \zeta \mathrm{Tr} \ \left(G^{\Lambda} \mathcal{Q}^{\Lambda} \tilde{\mathcal{V}}^{1,1}_{\bar{\phi},\phi} (\Gamma^{\Lambda}, G^{\Lambda}) \right) \ , \tag{59}$$

with

$$\tilde{\mathcal{V}}_{\bar{\phi},\phi} \left(\Gamma^{\Lambda}, G^{\Lambda} \right) = \left[\mathbf{1} - \begin{pmatrix} \zeta G^{\Lambda} & 0 \\ 0 & \left[G^{\Lambda} \right]^t \end{pmatrix} \begin{pmatrix} \mathcal{U}_{\bar{\phi},\phi} & \frac{\delta^2 \Gamma^{\Lambda}}{\delta \bar{\phi} \delta \bar{\phi}} \\ \frac{\delta^2 \Gamma^{\Lambda}}{\delta \phi \delta \phi} & \zeta \mathcal{U}^t_{\bar{\phi},\phi} \end{pmatrix} \right]^{-1} . \tag{60}$$

For later applications it is important to note that $\mathcal{U}_{\bar{\phi},\phi}$ as well as $\frac{\delta^2 \Gamma^{\Lambda}}{\delta \bar{\phi} \delta \bar{\phi}}$ and $\frac{\delta^2 \Gamma^{\Lambda}}{\delta \phi \delta \phi}$ are at least quadratic in the external sources. The initial condition for the exact partial functional differential equation (59) can either be obtained by lengthy but straightforward algebra not presented here, or by the following simple argument: at $\Lambda = \Lambda_{\mathrm{i}}$, $G^{0,\Lambda_{\mathrm{i}}} = 0$ (no degrees of freedom are "turned on") and in a perturbative expansion of the $\gamma_m^{\Lambda_{\mathrm{i}}}$ the only term which does not vanish is the bare two-particle vertex. We thus find

$$\Gamma^{\Lambda_{\mathrm{i}}} \left(\{\bar{\phi}\}, \{\phi\} \right) = A_{\mathrm{int}} \left(\{\bar{\phi}\}, \{\phi\} \right) \ . \tag{61}$$

An exact infinite hierarchy of flow equations for the γ_m^{Λ} can be obtained by expanding Eq. (60) in a geometric series and Γ^{Λ} in the external sources

$$\Gamma^{\Lambda} \left(\{\bar{\phi}\}, \{\phi\} \right) = \sum_{m=0}^{\infty} \frac{\zeta^m}{(m!)^2} \sum_{k'_1,\ldots,k'_m} \sum_{k_1,\ldots,k_m} \gamma_m^{\Lambda} \left(k'_1, \ldots, k'_m; k_1, \ldots, k_m \right) \bar{\phi}_{k'_1} \ldots \bar{\phi}_{k'_m} \phi_{k_m} \ldots \phi_{k_1} \ .$$

The equation for γ_0^{Λ} reads

$$\frac{d}{d\Lambda} \gamma_0^{\Lambda} = -\zeta \ \mathrm{Tr} \ \left(\mathcal{Q}^{\Lambda} \mathcal{G}^{0,\Lambda} \right) + \zeta \ \mathrm{Tr} \ \left(\mathcal{Q}^{\Lambda} \mathcal{G}^{\Lambda} \right) \ . \tag{62}$$

Via G^{Λ} the derivative of γ_0^{Λ} couples to the one-particle self-energy. For the flow of the self-energy we obtain

$$\frac{d}{d\Lambda} \gamma_1^{\Lambda}(k'; k) = \zeta \frac{d}{d\Lambda} \Sigma^{\Lambda}_{k',k} = \ \mathrm{Tr} \ \left(\mathcal{S}^{\Lambda} \gamma_2^{\Lambda}(k', \ldots; k, \ldots) \right) \ , \tag{63}$$

with the single scale propagator

$$\mathcal{S}^{\Lambda} = G^{\Lambda} \mathcal{Q}^{\Lambda} G^{\Lambda} \ . \tag{64}$$

Here $\gamma_2^{\Lambda}(k', \ldots; k, \ldots)$ is a matrix in the variables not explicitly written, i.e.

$$\left[\gamma_2^{\Lambda}(k', \ldots; k, \ldots) \right]_{q',q} = \gamma_2^{\Lambda}(k', q'; k, q).$$

The diagrammatic representation of Eq. (63) is of the form given in Fig. 2 supplemented by arrows at the lines and labels k and k'. The flow equation for γ_2^Λ reads

$$\begin{aligned}\frac{d}{d\Lambda}\gamma_2^\Lambda(k_1',k_2';k_1,k_2) &= \mathrm{Tr}\,\left(\mathcal{S}^\Lambda\gamma_3^\Lambda(k_1',k_2',\ldots;k_1,k_2,\ldots)\right)\\ &+\zeta\,\mathrm{Tr}\left(\mathcal{S}^\Lambda\gamma_2^\Lambda(\ldots,\ldots;k_1,k_2)\left[\mathcal{G}^\Lambda\right]^t\gamma_2^\Lambda(k_1',k_2';\ldots,\ldots)\right)\\ &+\zeta\,\mathrm{Tr}\left(\mathcal{S}^\Lambda\gamma_2^\Lambda(k_1',\ldots;k_1,\ldots)\mathcal{G}^\Lambda\gamma_2^\Lambda(k_2',\ldots;k_2,\ldots)\right.\\ &\left.+\zeta\,[k_1'\leftrightarrow k_2']+\zeta\,[k_1\leftrightarrow k_2]+[k_1'\leftrightarrow k_2',k_1\leftrightarrow k_2]\right)\,. \end{aligned} \tag{65}$$

A simplified (no arrows and labels) diagrammatic representation is shown in Fig. 2. Crucially, the structure of the funRG flow equations for a general quantum many-body problem is the same as the one obtained for the classical anharmonic oscillator discussed in the last section. The right hand sides of the coupled differential equations contain the feedback of the scale dependent vertices on the flow of these. The initial conditions for the $\gamma_m^{\Lambda_\mathrm{i}}$ can be obtained from Eq. (61) and are given by

$$\gamma_2^{\Lambda_\mathrm{i}}(k_1',k_2';k_1,k_2)=\bar{v}_{k_1',k_2',k_1,k_2}\quad,\quad\gamma_m^{\Lambda_\mathrm{i}}=0\ \text{ for }\ m\neq 2\,. \tag{66}$$

In contrast to our discussion of the above toy model we here refrain from explicitly presenting flow equations for γ_m^Λ with $m\geq 3$ since later on the set of differential equations is truncated by setting $\gamma_3^\Lambda=\gamma_3^{\Lambda_\mathrm{i}}=0$, which implies that vertices with $m\geq 3$ do not contribute.
The truncation procedure to obtain a finite, closed set of coupled equations is exactly the same as introduced for the toy model: for $m_\mathrm{c}\geq 2$, the vertex $\gamma_{m_\mathrm{c}+1}^\Lambda$ on the right hand side of the coupled flow equations is replaced by its initial condition $\gamma_{m_\mathrm{c}+1}^{\Lambda_\mathrm{i}}=0$. This set of differential equations can then be integrated over Λ starting at Λ_i down to Λ_f providing approximate expressions for the γ_m of the original (cutoff free) problem with $m\leq m_\mathrm{c}$. Expanding γ_m^Λ in terms of the bare interaction, conventional perturbation theory for the grand canonical potential, the self-energy, the effective interaction and higher order vertex functions can be recovered from an iterative treatment of the flow equations (see above). The vertex functions obtained from the truncated equations are at least correct up to order m_c in the bare interaction but contain (incomplete) resummations of higher order terms.

4 The quantum anharmonic oscillator

We now consider a second toy model namely the quantum harmonic oscillator with quartic perturbation. Due to the quantum nature the variables and external sources will become functions of the Matsubara frequency but do not carry single-particle quantum numbers. The funRG approach to this toy problem is in many respects as complex as the approach to problems of quantum dots with local Coulomb interaction coupled to noninteracting reservoirs (see below). In appropriate units the Hamiltonian is given by

$$\hat{H}=\frac{1}{2}\hat{p}^2+\frac{1}{2}\hat{x}^2+\frac{g}{4!}\hat{x}^4=\hat{H}_0+\hat{H}_\mathrm{int} \tag{67}$$

with the position operator $\hat{x}$, the momentum operator $\hat{p}$, and the coupling constant g. We here focus on $T=0$ and are interested in low-lying eigenenergies E_n as well as the (imaginary) time-ordered propagator

$$\mathcal{G}(\tau)=\langle E_0|\,\mathcal{T}\left[\hat{x}(\tau)\hat{x}(0)\right]|E_0\rangle\ , \tag{68}$$

with $|E_n\rangle$ being the eigenstates of Eq. (67) and $\hat{x}(\tau) = \exp(i\hat{H}\tau)\hat{x}\exp(-i\hat{H}\tau)$. We note that the Green function/propagator considered here differs structurally from the ones defined in the last section. The latter are expectation values of a product of a raising and a lowering operator, while the current one includes a product of the position operator which is the sum of two ladder operators: $\hat{x} = (\hat{a} + \hat{a}^\dagger)/\sqrt{2}$. The Fourier transform of this Green function can be written as

$$\mathcal{G}(i\omega) = \int_{-\infty}^{\infty} d\tau\, e^{i\omega\tau}\mathcal{G}(\tau) = \frac{1}{[\mathcal{G}^0(i\omega)]^{-1} - \Sigma(i\omega)}, \tag{69}$$

where we have introduced the self-energy Σ and the noninteracting propagator

$$\mathcal{G}^0(i\omega) = \frac{1}{\omega^2 + 1}. \tag{70}$$

In contrast to the more general notation used in the last section, propagators and the self-energy only depend on a single frequency and do no longer contain the energy conserving δ-function here. The propagator has the Lehmann representation [1, 2]

$$\mathcal{G}(i\omega) = \frac{1}{2}\sum_{n=1}^{\infty}\left[\frac{1}{i\omega + (E_n - E_0)} - \frac{1}{i\omega - (E_n - E_0)}\right]\left|\langle E_0|\,(\hat{a} + \hat{a}^\dagger)\,|E_n\rangle\right|^2. \tag{71}$$

The spectral weights

$$s_n = \left|\langle E_0|\,(a + a^\dagger)\,|E_n\rangle\right|^2$$

and energies E_n fulfill the f-sum rule [1]

$$1 = \sum_{n=1}^{\infty}(E_n - E_0)\; s_n. \tag{72}$$

It turns out that for coupling constants $g \leq 50$ considered here the sums in Eqs. (71) and (72) are dominated by the first few terms. For this reason only the first few eigenstates and eigenenergies are required to obtain accurate ("numerically exact") results for $\mathcal{G}(i\omega)$. These can quite easily be obtained by expressing $\hat{H}$ in the basis of eigenstates $|n\rangle$ of the unperturbed harmonic oscillator and numerically diagonalizing the upper left corner of the (infinite) matrix $\langle n|\,H\,|n'\rangle$ with $n, n' \leq n_\mathrm{c}$ and a sufficiently large n_c. For $g \leq 50$, $n_\mathrm{c} = 100$ turns out to be large enough to fulfill the sum rule Eq. (72) to very high precision.

4.1 Second order perturbation theory and mean-field theory

For the g-dependent part of the ground state energy second order perturbation theory yields

$$e_0^{(2)} = E_0^{(2)} - E_0^0 = \frac{1}{32}\,g - \frac{7}{1536}\,g^2 \tag{73}$$

and for the self-energy one obtains

$$\Sigma^{(2)}(i\omega) = -\frac{1}{4}\,g + \frac{1}{32}\,g^2 + \frac{1}{8}\,g^2\frac{1}{\omega^2 + 9}\,. \tag{74}$$

Within the funRG approximate expressions for

$$E_{n,0} = E_n - E_0$$

and s_n can be obtained from the poles and residues of the propagator $\mathcal{G}(i\omega)$. Furthermore, since Eqs. (71) and (72) are dominated by the first few terms we only consider $E_{1,0}$ and s_1. Second order approximations for these quantities are given by the smallest pole of $\mathcal{G}^{(2)}(i\omega) = \left[\omega^2 + 1 - \Sigma^{(2)}(i\omega)\right]^{-1}$ and the related residue. It is important to note that this approximation for $E_{1,0}^{(2)}$ agrees with $E_1^{(2)} - E_0^{(2)}$, where $E_1^{(2)}$ is determined directly from Rayleigh-Schrödinger perturbation theory, only up to second order in g, but is closer to the exact $E_{1,0}$.
Within mean-field theory, e_0^{MF} and a frequency independent Σ^{MF} are given by

$$e_0^{\mathrm{MF}} = \frac{1}{2}\sqrt{1 + \frac{g}{2}\left\langle X^2\right\rangle_{\mathrm{MF}}} - \frac{g}{8}\left\langle X^2\right\rangle_{\mathrm{MF}}^2 - \frac{1}{2} \quad , \quad \Sigma^{\mathrm{MF}} = -\frac{g}{2}\left\langle X^2\right\rangle_{\mathrm{MF}}$$

and $\left\langle X^2\right\rangle_{\mathrm{MF}}$ is the solution of the self-consistency equation

$$\left\langle X^2\right\rangle_{\mathrm{MF}} = \frac{1}{2}\,\frac{1}{\sqrt{1 + g\left\langle X^2\right\rangle_{\mathrm{MF}}/2}}.$$

From the mean-field propagator one obtains

$$E_{1,0}^{\mathrm{MF}} = \sqrt{1 - \Sigma^{\mathrm{MF}}} \quad , \quad s_1^{\mathrm{MF}} = \frac{1}{2}\,\frac{1}{\sqrt{1 - \Sigma^{\mathrm{MF}}}}.$$

4.2 The funRG approach

The grand canonical partition function of the Hamiltonian Eq. (67) can be written in terms of a Feynman path integral which corresponds to the functional integral presentation of the last section in position states instead of coherent ones. It reads

$$\frac{\mathcal{Z}}{\mathcal{Z}_0} = \frac{1}{\mathcal{Z}_0}\int \mathcal{D}\,\bar{x}\,x \exp\left\{(\bar{x}, \left[\mathcal{G}^0\right]^{-1} x)/2 - A_{\mathrm{int}}\left(\{\bar{x}\},\{x\}\right)\right\}, \tag{75}$$

with the interacting part of the action

$$A_{\mathrm{int}}\left(\{\bar{x}\},\{x\}\right) = \frac{g}{\beta\,4!}\sum_{n_1,\ldots,n_4}\delta_{n_1+n_2,n_3+n_4}\;\bar{x}(i\omega_1)\bar{x}(i\omega_2)x(i\omega_3)x(i\omega_4)\;, \tag{76}$$

bosonic Matsubara frequencies $\omega_j = 2\pi\, n_j/\beta$, and complex fields $\bar{x}(i\omega) = x(-i\omega)$.
As outlined in the last section and using the frequency cutoff Eq. (53) flow equations for the γ_m^Λ can be obtained. Here we focus on the equations in truncation order $m_{\mathrm{c}} = 2$. For $T \to 0$ and after introducing

$$e_0^\Lambda = \lim_{T\to 0} T\gamma_0^\Lambda$$

we find

$$\frac{d}{d\Lambda}e_0^\Lambda = -\frac{1}{2\pi}\ln\left[1 - \mathcal{G}^0(i\Lambda)\,\Sigma^\Lambda(\Lambda)\right]\;, \tag{77}$$

with the initial condition $e_0^{\Lambda_{\mathrm{i}}=\infty} = 0$. At the end of the flow, $e_0^{\Lambda_{\mathrm{f}}=0}$ directly provides the funRG approximation e_0^{funRG} for the g-dependent part of the ground state energy. The flow equation for the self-energy follows as

$$\frac{d}{d\Lambda}\Sigma^\Lambda(i\omega) = \frac{1}{2\pi}\,\frac{1}{\Lambda^2 + 1 - \Sigma^\Lambda(i\Lambda)}\;g^\Lambda(i\omega, -i\omega, i\Lambda, -i\Lambda)\;, \tag{78}$$

with initial condition $\Sigma^{\Lambda_i=\infty} = 0$. The funRG approximation for the self-energy is $\Sigma^{\rm funRG}(i\omega) = \Sigma^{\Lambda_i=0}(i\omega)$. Here g^Λ denotes the totally symmetric two-particle vertex which, in contrast to the vertex γ_2^Λ introduced in the last section, does not contain an energy conserving δ-function and factors of β. Due to energy conservation it depends only on three frequencies, but the fourth will nevertheless always be included in the following. To derive Eqs. (77) and (78) one has to deal with products of delta functions $\delta(|\omega| - \Lambda)$ and terms involving step functions $\Theta(|\omega| - \Lambda)$. These seemingly ambiguous expressions are well defined and unique if the sharp cutoff is implemented as a limit of increasingly sharp broadened cutoff functions Θ_ϵ, with the broadening parameter ϵ tending to zero. The expressions can then be conveniently evaluated by using the following relation, valid for arbitrary continuous functions f [10]:

$$\delta_\epsilon(x-\Lambda)\, f[\Theta_\epsilon(x-\Lambda)] \to \delta(x-\Lambda) \int_0^1 f(t)\, dt \;, \tag{79}$$

where $\delta_\epsilon = -d\Theta_\epsilon/d\epsilon$.
For g^Λ the flow equation reads

$$\begin{aligned}\frac{d}{d\Lambda} g^\Lambda(i\omega_1, i\omega_2, i\omega_3, -i\omega_1 - i\omega_2 - i\omega_3) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\,\nu \,\big[\mathcal{S}^\Lambda(i\nu)\mathcal{G}^\Lambda(i\nu - i\omega_1 - i\omega_2) \\ \times g^\Lambda(i\omega_1, i\omega_2, -i\nu, i\nu - i\omega_1 - i\omega_2) g^\Lambda(i\omega_3, -i\omega_1 - i\omega_2 - i\omega_3, -i\nu + i\omega_1 + i\omega_2, i\nu) \\ + (\omega_2 \leftrightarrow \omega_3) + (\omega_2 \leftrightarrow -\omega_1 - \omega_2 - \omega_3)\big] \,.\end{aligned} \tag{80}$$

As the single-scale propagator $\mathcal{S}^\Lambda(i\nu)$ contains a factor $\delta(|\nu| - \Lambda)$ the integral over ν in Eq. (80) can be performed analytically.
To numerically solve the set of differential equations (77), (78), and (80) we have discretized the frequencies (which at $T = 0$ are continuous) on a linear mesh $\omega_j = j\delta$ with $j = -j_0, -j_0 + 1, \ldots, j_0$. By increasing j_0 and decreasing δ convergence can be achieved up to the required accuracy. For our purposes $j_0 = 40$ and $\delta = 0.5$ turned out to be appropriate. This leads to a set of roughly 5.3×10^5 coupled equations. The Λ-integration is started at $\Lambda_i = 10^5$ making sure that further increasing Λ_i does not lead to significant changes in the results. Figure 3 shows comparisons of e_0, $E_{1,0}$, and s_1 for the different approximations considered here and the exact results. Although the approximate funRG correctly reproduces only the first two derivatives with respect to g at $g = 0$, it gives very accurate results even up to $g = 50$, while conventional second order perturbation theory can only be trusted for $g < 1$. This provides yet another impressive example of the power of "renormalization group enhanced perturbation theory". To avoid the problem of analytic continuation from imaginary to real frequencies (see lecture B4 by S. Wessel) the results for $e_{1,0}^{\rm funRG}$ and $s_1^{\rm funRG}$ were obtained by fitting a function $a/(\omega^2 + b^2)$ with a and b as fitting parameters to $\mathcal{G}^{\rm funRG}(i\omega) = \left[\omega^2 + 1 - \Sigma^{\rm funRG}(i\omega)\right]^{-1}$. Assuming this fitting form we have used that the spectral function is dominated by the first peak.[4] For the problem studied also mean-field theory leads to fairly accurate results (but not as good as the funRG). This is related to the fact that low-lying eigenstates of the Hamiltonian in Eq. (67) can be described quite well by the eigenstates of a harmonic oscillator with a modified frequency determined self-consistently.

[4]For symmetry reasons the s_n with even n vanish. For the g considered here the exact spectral weight s_3 is roughly three orders of magnitude smaller than s_1. Thus extracting more than $s_1^{\rm funRG}$ and $E_{1,0}^{\rm funRG}$ is beyond the accuracy of our numerical treatment of the flow equations.

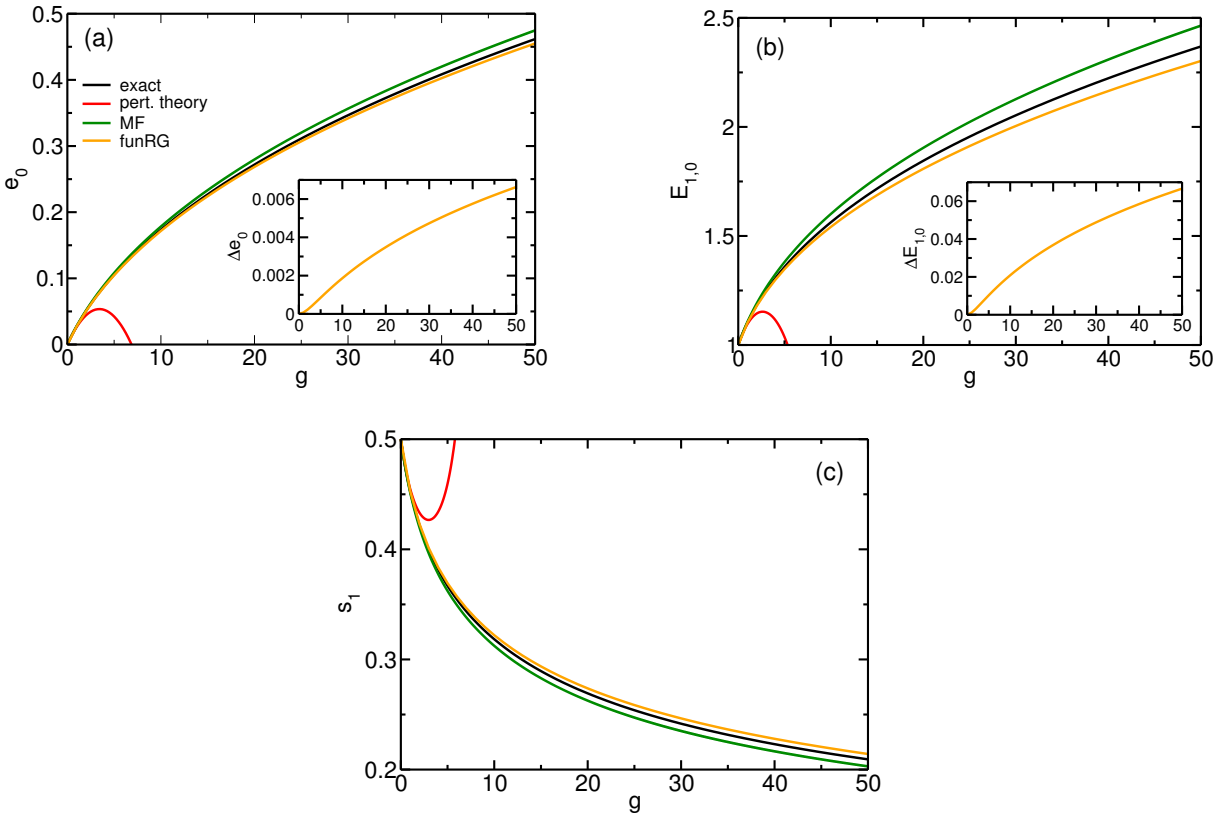


Fig. 3: *(a) Coupling constant dependent part of the ground state energy $e_0 = E_0 - E_0^0$ as a function of g. The different approximations are compared to the exact result. The inset shows the difference between the exact result and the funRG approximation. (b) Energy difference $E_{1,0}$ of the first excited state and the ground state. (c) The spectral weight s_1 of the first peak.*

5 The single impurity Anderson model

We will consider two models of quantum dots with local correlations coupled to two noninteracting leads (reservoirs). The single impurity Anderson model (SIAM) is the prototype model to describe quantum dots with level spacings Δ much larger than any other energy scale of the problem, e.g. the level-lead coupling Γ or the charging energy U. Considering $T \ll \Delta$ one can then focus on a single dot level and the Hamiltonian consists of three parts (see also the lecture B3 by T. Costi)

$$\hat{H} = \hat{H}_{\rm res} + \hat{H}_{\rm dot} + \hat{H}_{\rm coup}. \tag{81}$$

The left (index L) and right (index R) reservoirs are described by noninteracting electrons with (for simplicity equal) single-particle dispersion $\epsilon_{\vec{k}}$

$$\hat{H}_{\rm res} = \sum_{\vec{k}\sigma l} \epsilon_{\vec{k}} \hat{c}^\dagger_{\vec{k}\sigma l} \hat{c}_{\vec{k}\sigma l}, \tag{82}$$

where $\hat{c}_{\vec{k}\sigma l}$ denote fermionic annihilation operators for electrons with momentum $\vec{k}$ and spin direction $\sigma = \uparrow, \downarrow (= \pm)$ in lead $l = L/R$. The impurity Hamiltonian contains both single- and two-particle terms,

$$\hat{H}_{\rm dot} = \sum_\sigma \left(\epsilon + \sigma \frac{B}{2}\right) \hat{d}^\dagger_\sigma \hat{d}_\sigma + U \left(\hat{d}^\dagger_\uparrow \hat{d}_\uparrow - \frac{1}{2}\right)\left(\hat{d}^\dagger_\downarrow \hat{d}_\downarrow - \frac{1}{2}\right), \tag{83}$$

with $\hat{d}_\sigma$ annihilating an electron of spin σ located at the impurity. The single-particle energy ϵ was shifted such that at $B = 0$, $\epsilon = 0$ corresponds to the point of particle-hole symmetry, and U and B denote the strength of the local Coulomb repulsion and of the local Zeeman field, respectively. Finally, the coupling between the dot and the reservoirs is given by

$$\hat{H}_{\text{coup}} = -t \sum_{\sigma l} \hat{c}^\dagger_{\sigma l} \hat{d}_\sigma + \text{H.c.} \,, \tag{84}$$

where $\hat{c}_{\sigma l}$ is the local electron operator $\hat{c}_{\sigma l} = \sum_{\vec{k}} \hat{c}_{\vec{k}\sigma l}/\sqrt{N}$ with N the number of lead lattice sites. For simplicity we assume that the tunneling from the dot to the left and right reservoirs is equal. This assumption can easily be relaxed. In lecture B3 T. Costi already gave you an overview of the physics of this model with a particular focus on the Kondo effect.
In order to apply the funRG scheme presented above to the SIAM Hamiltonian Eq. (81), the noninteracting reservoirs have to be integrated out. This leads to the reservoir self-energy as introduced by S. Jakobs in lecture B2. Thereafter, instead of dealing with an infinite system (after taking $N \to \infty$) one only needs to consider two interacting (spin up and down) particles. For the noninteracting dot Green function we obtain

$$G^0_\sigma(i\omega) = \frac{1}{i\omega - \epsilon - \sigma B/2 + i\,\text{sgn}(\omega)\Gamma} \,, \tag{85}$$

where the hybridization $\Gamma = \Gamma_L + \Gamma_R = 2\pi|t|^2 \sum_{\vec{k}} \delta(E - \epsilon_{\vec{k}})/N$ is assumed to be energy-independent (wide-band limit).
Here we will primarily consider the case $T = 0$ in which correlations are most prominent. Perturbation theory of this model in the local two-particle interaction U is regular (the different terms do not diverge) but does not show Kondo physics [1]. The mean-field solution of the SIAM is e.g. discussed in the text book Ref. [3]. In particular, for $\epsilon = 0$ and $B = 0$ it shows spontaneous spin symmetry breaking if $U/\Gamma > \pi$, that is when entering the regime in which Kondo physics starts to matter. This is an artifact of the approximation and provides sufficient motivation to use alternative resummation schemes such as funRG.

5.1 The funRG to truncation order $m_{\text{c}} = 1$

We will start out by considering funRG in the simplest truncation scheme with $m_{\text{c}} = 1$. Within this the two-particle vertex does not renormalize, that is it remains the bare interaction, and the interacting part of the self-energy is frequency independent. This directly follows from the Hartree-Fock-like topology of the diagrammatic representation of $\dot{\gamma}^\Lambda_1$ and the frequency independence of the bare two-particle interaction. We use the frequency cutoff Eq. (53). The set of flow equations derived within the general formalism thus boils down to two equations for the dot self-energies of the up and down spin directions. Combined with the unrenormalized (bare) level positions $\epsilon + \sigma B/2$ we end up with flow equations for effective renormalized ones $V^\Lambda_\sigma = \epsilon + \sigma B/2 + \Sigma^\Lambda_\sigma$ which read

$$\frac{d}{d\Lambda} V^\Lambda_\sigma = \frac{U V^\Lambda_{-\sigma}/\pi}{(\Lambda + \Gamma)^2 + (V^\Lambda_{-\sigma})^2}, \tag{86}$$

with initial conditions $V^{\Lambda=\infty}_\sigma = \epsilon + \sigma B/2$. At the end of the flow, the renormalized level position $V_\sigma = V^{\Lambda=0}_\sigma$ determines the dot spectral function

$$\rho_\sigma(\omega) = -\frac{1}{\pi} \text{Im}\, G_\sigma(\omega + i\eta) = \frac{1}{\pi} \frac{\Gamma}{(\epsilon - V_\sigma)^2 + \Gamma^2}, \tag{87}$$

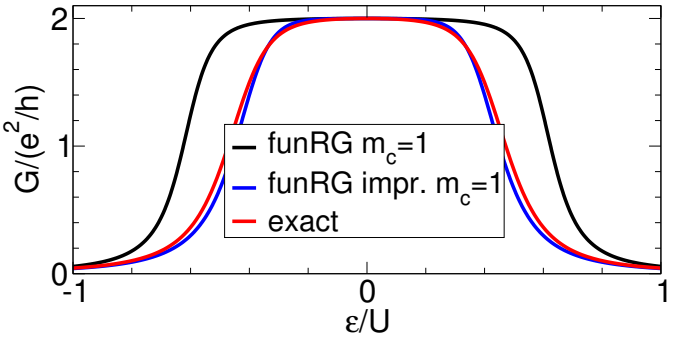


Fig. 4: *Level position dependence of the linear conductance for* $U/\Gamma = 4\pi$ *at* $B = 0$.

which is a Lorentzian of full width 2Γ and height $1/\pi\Gamma$ centered around V_σ. The approximation $G_\sigma(\omega + i\eta)$ of the single-particle Green function (the dot propagator) is given by Eq. (85) with the bare level position $\epsilon + \sigma B/2$ replaced by the renormalized one. We note that in the present truncation order the analytic continuation from Matsubara to real frequencies (see lecture B4 by S. Wessel) is straightforward as the self-energy is frequency independent. It is very appealing and helps a lot in developing an intuition of interaction effects, that in truncation order $m_c = 1$ at the end of the funRG flow one has to deal with an effective single-particle problem (frequency independent self-energy). The interaction effects are coded in the parameter dependence of the self-energy. This is reminiscent of the Hartree-Fock approximation with the crucial difference that for the SIAM funRG avoids the spurious spontaneous symmetry breaking.
Let us for now consider the case of vanishing Zeeman splitting. The solution of the differential Eq. (86) at $\Lambda = 0$ can be given in closed from as

$$\frac{vJ_1(v) - \gamma J_0(v)}{vY_1(v) - \gamma Y_0(v)} = \frac{J_0(e)}{Y_0(e)}, \tag{88}$$

with $v = V\pi/U$, $e = \epsilon\pi/U$, $\gamma = \Gamma\pi/U$, and Bessel functions J_n, Y_n. For $|\epsilon| < V_c$ this equation has a solution with small $|V|$, where $v_c = V_c\pi/U$ is the first zero of J_0 corresponding to $V_c \simeq 0.77\,U$. For $U \gg \Gamma$ the crossover to a solution with $|V|$ being of order U (for $|\epsilon| > V_c$) is fairly sharp. Expanding both sides of Eq. (88) for small $|v|$ and $|e|$ gives

$$V = \epsilon \exp\left(-\frac{U}{\pi\Gamma}\right). \tag{89}$$

The consequent exponential pinning of the spectral weight in Eq. (87) at the chemical potential for small $|\epsilon|$ and the sharp crossover to a V of order U when $|\epsilon| > V_c$ is characteristic of Kondo physics (see lecture B3 by T. Costi). The pinning is captured by our approximate approach despite the fact that the true spectral function has a very different form than the Lorentzian Eq. (87) with a very sharp Kondo resonance of width $T_K \sim \exp[-\pi U/(8\Gamma)]$ [1, 3] and Hubbard satellites. The Kondo scale T_K is a prototype emergent energy scale as discussed in the introduction.
On the one hand side we must thus conclude that the present approximation scheme cannot be used when aiming at the dot spectral function. On the other hand it becomes meaningful when considering observables which are given by the frequency integrated spectral function such as the dots occupancy (not shown here) or the spectral weight at the chemical potential $\rho(\omega = 0)$ such as the linear response conductance (see lectures A9 by S. Tsukamoto, B2 by S. Jakobs, B3 by T. Costi, B5 by H. Schoeller, and C7 by M. Wegewijs). In general the linear conductance is

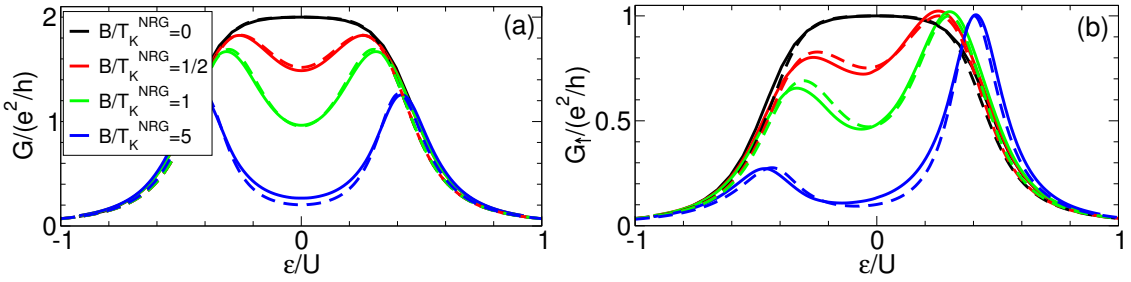


Fig. 5: *Level position dependence of the total conductance G (a) and the spin-resolved one (b) of a single dot with $U/\Gamma = 3\pi$ for different Zeeman fields B expressed in units of the $\epsilon = 0$ Kondo temperature $T_{\rm K}^{\rm NRG}/\Gamma = 0.116$. Solid lines: NRG data from Ref. [11]. Dashed line: funRG approximation with flow of static part of the vertex.*

given by the current-current correlation function (Kubo formula) [1, 3]. For the present model at zero temperature the exact conductance has the simple form[5]

$$G(\epsilon) = G_\uparrow(\epsilon) + G_\downarrow(\epsilon), \quad G_\sigma(\epsilon) = \frac{e^2}{h}\,\pi\Gamma\rho_\sigma(0) \tag{90}$$

in terms of the dot spectral function. The pinning of the spectral weight when varying the bare level position ϵ then leads to a plateau like resonance in the linear conductance $G(\epsilon)$ which has a completely different line shape as the $U = 0$ one which is Lorentzian. This is shown for rather strong interaction $U/\Gamma = 4\pi$ in Fig. 4. For the SIAM the $T = 0$ conductance can be computed exactly using a method called Bethe ansatz. Obviously to order $m_{\rm c} = 1$ the ϵ regime of the pinning is overestimated. This weakness can be cured by slightly extending the $m_{\rm c} = 1$ scheme.

Flow of the static part of the two-particle vertex: The next step is to include the flow of the two-particle vertex γ_2^Λ, but still neglect the three-particle vertex γ_3^Λ as well as the frequency dependence of γ_2^Λ. For the effective flowing local two-particle interaction we obtain the equation

$$\frac{d}{d\Lambda}U^\Lambda = \frac{2\,\left(U^\Lambda\right)^2\,V_\uparrow^\Lambda\,V_\downarrow^\Lambda/\pi}{\left[(\Lambda+\Gamma)^2+(V_\uparrow^\Lambda)^2\right]\left[(\Lambda+\Gamma)^2+(V_\downarrow^\Lambda)^2\right]}, \tag{91}$$

with initial condition $U^{\Lambda=\infty} = U$. It complements Eq. (86) for the flow of the level positions in which on the right hand sides U must be replaced by the flowing interaction U^Λ. The set of coupled differential equations can easily be solved on a computer using standard algorithms. As shown in Fig. 4 keeping the flow of the static part of the interaction improves the quality of the approximation significantly. We emphasize, that we cannot speak of this approximation as the truncation order $m_{\rm c} = 2$ scheme as the later requires to keep the dependence of the two-particle vertex on three frequencies which then also leads to a frequency dependent self-energy.

Zeeman field dependence of the linear conductance: We next consider the case of finite Zeeman fields. For $B > 0$ the Kondo resonance in the NRG solution of the spectral function splits into two peaks with a dip at $\omega = 0$, resulting in a dip of $G(\epsilon)$ at $\epsilon = 0$. In Fig. 5 we compare the total $G = G_\uparrow + G_\downarrow$ and partial $G_\uparrow$ conductance obtained from the above funRG truncation scheme including the flow of the effective interaction with NRG results (see lecture

[5]Note that we use the same letter for the linear conductance and the Green function. From the context it should become clear which of the two is meant.

B3 by T. Costi) for different B expressed in units of $T_{\rm K}^{\rm NRG} = 0.116\Gamma$, where $T_{\rm K}^{\rm NRG}$ is the width of the Kondo resonance at the particle-hole symmetric point $\epsilon = 0$ obtained by NRG [11]. As the data for $B/T_{\rm K}^{\rm NRG} = 1$ suggest (see Fig. 5 (a)) the Zeeman field dependence of the conductance allows to extract the Kondo scale; the latter is given by the B at which the conductance is suppressed by a factor $1/2$ compared to the unitary value $2e^2/h$. In fact, at particle-hole symmetry $T_{\rm K}^{\rm funRG} \sim \exp\left[-U/(\pi 8)\right]$. The funRG approach to order $m_{\rm c} = 1$ thus leads to a resummation which implies an exponential dependence on U/Γ; the prefactor of U/Γ in the argument is $1/\pi$ which is rather close to the exact prefactor $\pi/8$.

Applications to more complex dot models: As the physics discussed in the present section was known prior to the application of funRG to truncation order $m_{\rm c} = 1$ the above results can merely be seen as a verification ("benchmarking") that this approximation provides a reasonable starting point to study more complex dot models, such as e.g. multi-level dots. The effort to numerically solve the funRG flow equations only grows moderately with increasing number of correlated degrees of freedom which must be contrasted to other approaches in which the computational limits are reached more quickly. The present funRG approximation was e.g. used to study two-level dots showing novel resonance [12] (with experimental verification in Ref. [13]), to investigate the so-called phase lapse problem of mesoscopic physics experimentally uncovered in the mid nineties [14] in which several dot levels matter [15], and the Josephson current through a single-level quantum dot with superconduction leads [16] also investigated experimentally [17]. We reemphasize, that due to the lack of frequency dependence of the self-energy the $m_{\rm c} = 1$ truncation should not be used when aiming at dynamical properties or finite temperature effects. To tackle those one has to resort at least to the order $m_{\rm c} = 2$ truncation at which the effective interaction and the self-energy become frequency dependent. We next discuss the latter in its application to the SIAM.

5.2 The funRG to truncation order $m_{\rm c} = 2$

After integrating out the leads in the SIAM one has to deal with two interacting degrees of freedom (spin-up or -down electron on the dot level). The complexity of setting up funRG flow equations to order $m_{\rm c} = 2$ is thus comparable to the one for the toy problem of the quantum harmonic oscillator with quartic perturbation discussed in Sect. 4.2.

We refrain from presenting the derivation of the corresponding set of coupled differential flow equations and here do not even give these equations. These and a detailed derivation can be found in Refs. [18] and [19]. We instead focus on a discussion of the merits and shortcomings. The numerical solution of the funRG flow equations and the computation of observables provides two challenges: (i) one has to ensure that the discretization of frequency space was chosen fine enough and (ii) that funRG provides imaginary frequency results. The required analytic continuation (see lecture B4 by S. Wessel) is an ill-posed problem and based on numerical data contains several pitfalls. Figure 6 shows a comparison of the dots spectral function at $T = \epsilon = B = 0$ obtained by a (stable) analytic continuation (by Padé approximation) of funRG data with highly accurate NRG results and other approximate ones. The interaction is $U/\Gamma = 5$ being at the boundary of the regime in which the Kondo effect starts to dominate the physics. This can e.g. seen by the sizable narrowing of the central resonance compared to the noninteracting width Γ. The funRG spectral function nicely agrees to the NRG data and is clearly superior to second order perturbation theory (in particular at small frequencies; see the inset of Fig. 6). In addition, the outcome of the restricted Hartree-Fock approximation is shown. In the latter one considers the not spin symmetry broken solution of the self-consistency equation. It

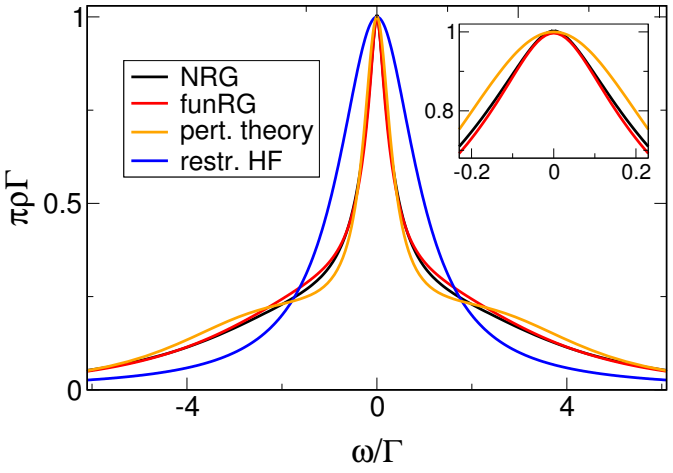


Fig. 6: *Dot spectral function at $T = \epsilon = B = 0$ for $U/\Gamma = 5$. The funRG results are compared to highly accurate NRG data, second order perturbation theory in U, and the restricted Hartree-Fock approximation.*

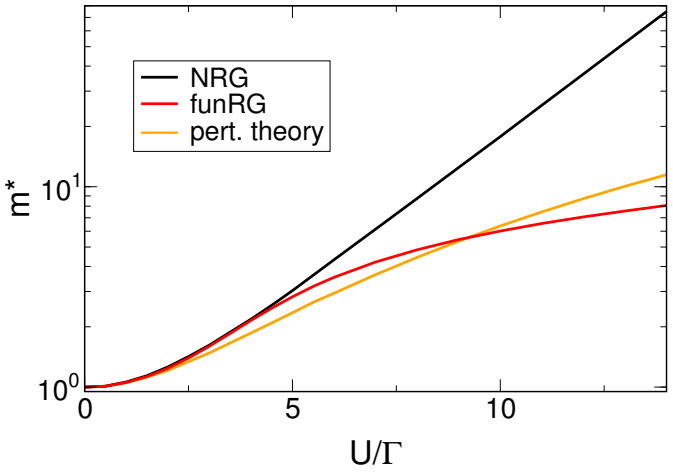


Fig. 7: *The renormalized mass as a function of U/Γ (for $B = 0$). The funRG results are compared to NRG and second order perturbation theory.*

does not capture the narrowing of the central peak.
Going to significantly larger U/Γ that is approaching the strong coupling regime is not possible within the funRG approach based on truncation schemes which are guided by perturbation theory (as considered here). Even if a stable analytic continuation can be achieved—the latter is getting more and more difficult with increasing U—the agreement with NRG data starts to deteriorate. In fact, at large U/Γ the width of the central peak which is proportional to the Kondo scale does not become exponential in U/Γ within funRG (to truncation order $m_c = 2$). A measure for the inverse of this width, namely the renormalized mass m^*, is shown in Fig. 7. While for $U/\Gamma \lessapprox 5$ the agreement with NRG data is excellent it quickly starts to become bad at larger U/Γ. The second order perturbation theory results start to deviate significantly earlier form the NRG curve. Fortunately, many of the experimental systems showing indications of Kondo physics have effective interactions of the order of those accessible to funRG. It thus becomes a useful tool despite its lack of capturing the strong coupling physics. Furthermore, using certain tricks computing the linear conductance does not require a full fledged analytic

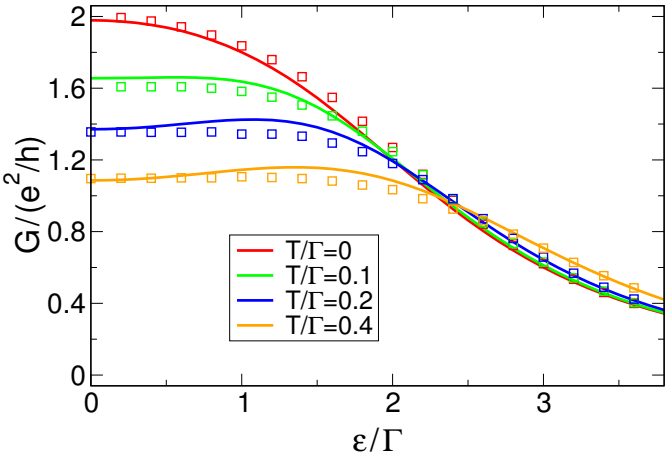


Fig. 8: *The linear conductance as a function of level position for different temperatures T at Zeeman field $B = 0$. Lines: NRG; symbols: funRG.*

continuation of numerical funRG data and can be obtained in a more stable way [20]. In Fig. 8 we compare the level position dependence of funRG results for the linear conductance G at different temperatures with NRG data for $U/\Gamma = 4$ (at $B = 0$).

Two alternative funRG based approaches to the SIAM which do not rely on the assumption that U constitutes a small parameter are described in Refs. [24] and [25]. Both have their individual merits and shortcomings.

By considering a functional integral representation of the generating functional for connected Green functions on the Keldysh contour (see lecture B2 by S. Jakobs) the present funRG scheme was extended to study the SIAM in a bias voltage driven nonequilibrium steady state; for the order $m_c = 1$ truncation, which has a rather limited range of applicability, see Ref. [21] while the superior $m_c = 2$ truncation is discussed in Ref. [22]. Note that in nonequilibrium it is advantageous to use a different cutoff scheme which in contrast to the sharp energy cutoff preserves causality [22]. The current-(bias-)voltage characteristics computed within the second scheme nicely agrees to those obtained by alternative numerical or analytical methods [23]. We emphasize that at finite bias voltage no exact results for the SIAM are known. The nonequilibrium steady-state physics of correlated quantum dots is a topic of great current interest.

6 The interacting resonant level model

Up to now we only applied funRG to models for which the perturbation theory in U is regular. We thus shed some light on the "RG enhanced perturbative" nature of funRG. In the SIAM we in addition exemplified if and how it deals with emergent many-body low-energy scales such as the Kondo scale T_K. We now proceed with a model which is also characterized by an emergent scale which strongly depends on the two-particle interaction, namely the interacting resonant level model (IRLM). In contrast to the SIAM in the IRLM the creation of the emergent scale is indicated by logarithmic terms appearing in perturbation theory.

The IRLM is a fundamental model to describe correlated quantum dots which are either spin-polarized or dominated by charge (instead of spin) fluctuations. Its Hamiltonian has the form Eq. (81) with the dot region consisting of three lattice sites which each can only be occupied by a single spinless fermion. The reservoir fermions are also assumed to be spinless. The onsite

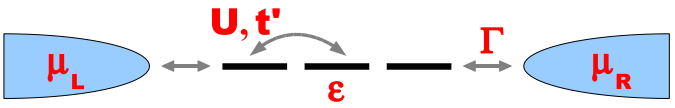


Fig. 9: *Sketch of the IRLM.*

energy ϵ of the central site can be varied by a gate voltage and the fermions on the dot interact via a density-density-type interaction ($\hat{n}_j = \hat{d}_j^\dagger \hat{d}_j$)

$$\hat{H}_{\rm dot} = \epsilon \hat{n}_2 - U\left(\frac{1}{2}\hat{n}_1 + \hat{n}_2 + \frac{1}{2}\hat{n}_3\right) + t'(\hat{d}_1^\dagger \hat{d}_2 + \hat{d}_2^\dagger \hat{d}_3 + \text{H.c.}) + U(\hat{n}_1\hat{n}_2 + \hat{n}_2\hat{n}_3). \tag{92}$$

The diagonal single particle terms in the second addend are chosen such that $\epsilon = 0$ again corresponds to the particle-hole symmetric point. The model is sketched in Fig. 9. Here we are mainly interested in the so-called scaling limit $t' \ll \Gamma$ in which the two dot sites 1 and 3 are effectively incorporated into the leads and the detailed structure of the leads does not matter (similar to the wide-band limit considered for the SIAM); the physics becomes universal.
We focus on the discussion of the results and refrain from giving the funRG flow equations in their full glory. They can be found in Refs. [26] (for equilibrium and the bias-voltage driven nonequilibrium steady state) and [27] (for the nonequilibrium relaxation dynamics).
We first consider equilibrium properties of the IRLM. Computing the Fock term of first order perturbation theory (the Hartree term vanishes) leads to an effective level-lead hopping

$$\left(\frac{t'^{\,\rm pert}}{t'}\right)^2 = 1 - \frac{4U}{\pi\Gamma}\,\ln\left(\frac{t'}{\Gamma}\right). \tag{93}$$

In the scaling limit $t' \gg \Gamma$ the logarithm grows and the first order correction is small compared to the bare value only for very small U/Γ. In this sense perturbation theory shows a logarithmic divergence. It must be contrasted to the type of divergencies appearing in the infrared limit of higher dimensional models (see the introduction) but still requires an appropriate regularization/resummation. In fact, the logarithmic term indicates the appearance of a renormalized level-lead coupling which scales as a power law with U-dependent exponent. The order $m_{\rm c} = 1$ funRG flow equation for the renormalized hopping at $\varepsilon = 0$ reads

$$\frac{d}{d\Lambda}t'^{\Lambda} = -\frac{U}{\pi\Gamma}\frac{t'^{\Lambda}/\Gamma}{(\Lambda/\Gamma)^2 + \Lambda/\Gamma + 2(t'^{\Lambda}/\Gamma)^2}, \quad t'^{\Lambda=\infty} = t'. \tag{94}$$

In the limit of small $U/(\pi\Gamma)$ and t'/Γ it can be solved analytically

$$\left(\frac{t'^{\Lambda=0}}{t'}\right)^2 \sim \left(\frac{t'}{\Gamma}\right)^{-\frac{4U}{\pi\Gamma}}, \tag{95}$$

which shows the above mentioned power-law scaling. This scale manifests in e.g. the line width of the linear conductance as a function of level position ϵ. In Fig. 10 order $m_{\rm c} = 1$ funRG results are compared to highly accurate results obtained by the numerical density-matrix renormalization group (DMRG) approach. With increasing U the width of the resonance increases in accordance with Eq. (95). The funRG provides an excellent approximation up to sizable U/Γ. To obtain funRG results for the linear conductance G is computationally rather cheap which must be contrasted to the heavy numerics required within DMRG.

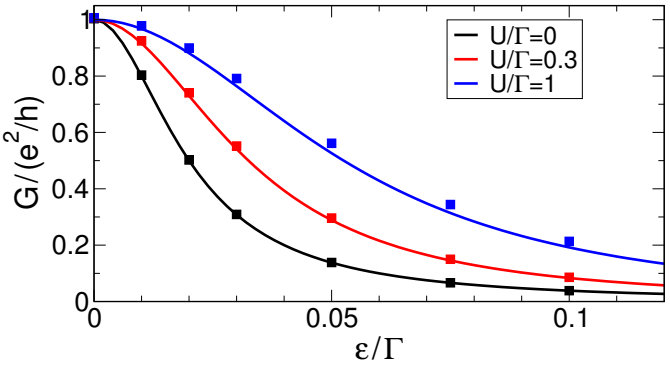


Fig. 10: *Linear conductance G as a function of level position ϵ for the IRLM at $T = 0$. Lines: DMRG from Ref. [28]; symbols: funRG.*

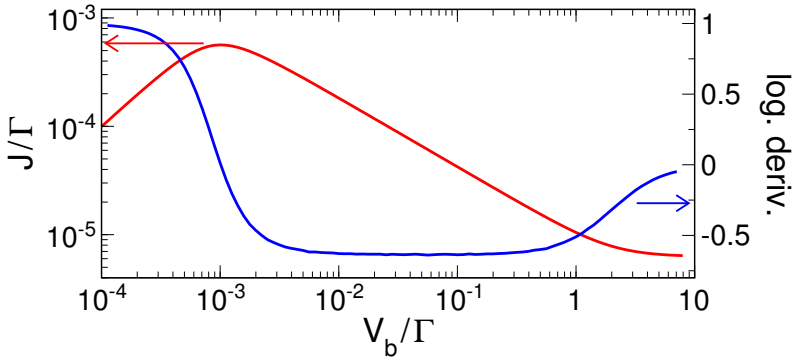


Fig. 11: *Zero temperature current-voltage characteristics of the IRLM for $U/\Gamma = 1$, $t'/\Gamma = 10^{-3}$, and $\epsilon = 0$ (in units such that $\hbar = 1 = e$).*

6.1 The bias-voltage driven steady state

After having gained confidence that funRG to truncation order $m_c = 1$ captures the renormalization of the level-lead coupling in equilibrium we proceed and use Keldysh funRG to study the bias voltage driven steady state. The bias voltage V_b is given by the difference of chemical potentials of the left and right reservoirs (see Fig. 9). After driving the system out of equilibrium the infinite reservoirs ensure that a steady-state is reached and the Keldysh funRG formalism can directly be set up for single time (or real frequency) Green functions (see lecture B2 by S. Jakobs). As mentioned above in nonequilibrium it is advantageous to use a different cutoff scheme than the sharp energy cutoff as the later spoils causality at order $m_c + 1$. The so-called reservoir cutoff in which a structureless auxiliary reservoir is coupled with strength Λ to every interacting site turned out to be convenient [22]. The cutoff Λ is then sent from infinity down to zero; at the end of the funRG flow the auxiliary reservoirs are decoupled and the system of interest is recovered.

In Fig. 11 the bias voltage dependence of the current J at particle-hole symmetry is shown (red curve, left y-axis). For V_b much smaller than the renormalized resonance width—the latter defines an emergent energy scale often denoted T_K in analogy to the Kondo scale of the SIAM—the current is ohmic $J \sim V_b$. This becomes clear from the blue curve (right y-axis scale) which shows the logarithmic derivative of the current. A plateau in the blue curve signals power-law scaling of J with the plateau value being the exponent; for $V_b \ll T_K$ a plateau with value 1 develops. For $T_K \ll V_b \ll \Gamma$ the current decreases with increasing V_b (negative differential

conductance). This second regime of power-law scaling is characterized by a U-dependent exponent which to leading order in U is given by $-2U/(\pi\Gamma)$ [29]. The rather flexible funRG approach now allows to investigate the current in due detail also considering $\epsilon \neq 0$, left-right asymmetric level-lead couplings [30], and finite temperatures [31]. It turns out that the current is affected by a subtle interplay of the different energy scales which can nicely be captured using funRG. In particular, it was shown that the bias voltage does not enter the current as a simple infrared cutoff.
In addition, funRG was used to investigate the linear response as well as nonequilibrium thermoelectric transport properties of the IRLM with a special focus on the output power and efficiency when using a IRLM-like quantum dot for converting waste heat into electrical energy [32]. A complementary approach very successfully used to study the nonequilibrium physics of the IRLM is the real-time RG [30, 33].

6.2 Quench dynamics

As a final application of funRG to study correlated quantum dots we discuss the nonequilibrium relaxation dynamics of the IRLM. A simple setup with a product initial state of the reservoirs each separately being in equilibrium at temperatures $T = 0$ and chemical potentials $\mu_L \neq \mu_R$ and the dot being empty was investigated using real-time RG in Ref. [30]. At time $t = 0$ the dot-lead coupling is turned on and fermions start to flow onto and through the dot. Within funRG one can reproduce the results of this study including the appearance of oscillations and of power-law corrections with U-dependent exponents to the exponential decay towards the steady state with interaction dependent decay rates [27]. To exemplify the flexibility of the funRG approach to the nonequilibrium real-time dynamics of correlated quantum dots we consider a more complex situation in which one prepares the system according to the above initial state, turns on Γ at $t = 0$, waits until a steady state is reached, and at time $t_\mathrm{q} > 0$ quenches the system by abruptly changing the on-site energy of the central dot site (see Fig. 9) from $\epsilon_\mathrm{ini} = 0$ to $\epsilon_\mathrm{fin} > 0$. For $t \gg t_\mathrm{q}$ the system relaxes towards a second steady state. Quenches of this type are currently heavily investigated in condensed matter physics.
To study the quench dynamics using Keldysh funRG one cannot resort to frequency space and has to consider two-time Green functions (see lecture B2 by S. Jakobs). The funRG flow equations to truncation order $m_\mathrm{c} = 1$ are given (and derived) in Refs. [27] and [34]. We choose $t'/\Gamma = 0.5$, which is no longer deep in the scaling limit. In this case one can resolve that the dot consists of three levels (see Fig. 9). Note that in the time evolution in general the currents across the left and right dot-lead bonds must not be the same. They become equal only at long times, that is in the steady state.
Figure 12 shows the time evolution of the occupancies of the three sites as well as the left current $J_L(t)$ for $V_\mathrm{b}/\Gamma = 2$ following from a numerical solution of the funRG flow equations. The occupancy of the middle site and the current drop quickly after the ϵ-quench at $\Gamma t_\mathrm{q} = 20$ as the onsite energy is raised far above the chemical potential. Whereas in the noninteracting case $n_2(t) = \langle \hat{n}_2(t) \rangle$ smoothly decreases (up to weakly developed shoulders; dashed line in the upper panel) this depletion is periodically interrupted for $U > 0$ (solid line; see the ellipses in the upper panel). This effect, which clearly goes beyond the renormalization of the decay rates by the local two-particle interaction, can be understood as follows. The interaction enhances the coherence of the system [30, 27] leading to an increase of the oscillation amplitudes. Plateau-like features show up for time regimes where maxima of the current leaving the left reservoir coincide with the decreasing part of the oscillatory occupancy of the first site $n_1(t)$. The combi-

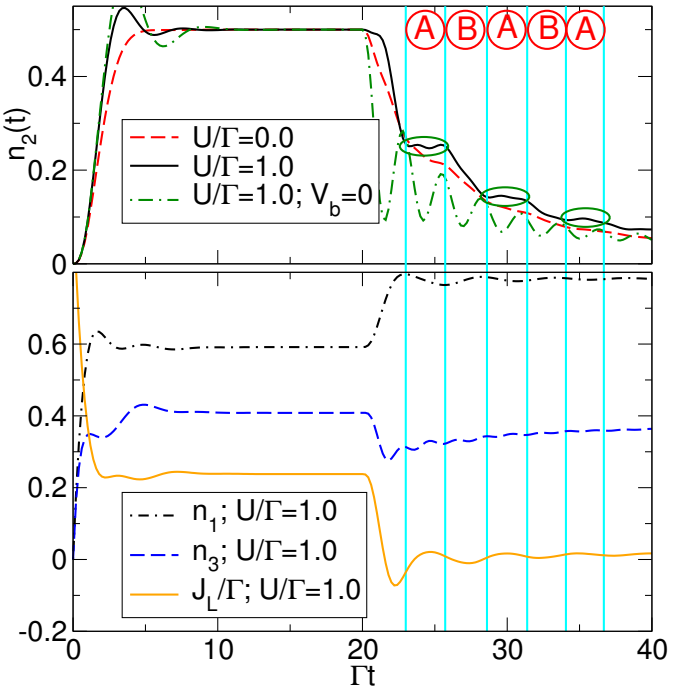


Fig. 12: *Top: Time evolution of the occupancy $n_2(t)$ of the central level as a function of time. The quench from $\epsilon_{\rm ini}/\Gamma = 0$ to $\epsilon_{\rm fin}/\Gamma = 2$ is performed at time $\Gamma t_{\rm q} = 20$. At this time the system has reached a steady state after initially being prepared according to the protocol described in the main text. The other parameters are $t'/\Gamma = 0.5$ and $V_{\rm b}/\Gamma = 2$. The plateau-like features interrupting the exponential decay beyond $\Gamma t_{\rm q} = 20$ are highlighted by ellipses. For comparison $n_2(t)$ for $U/\Gamma = 1$ and $V_{\rm b} = 0$ is shown as well. Bottom: Time evolution of $n_{1/3}(t)$ and the left current $J_L(t)$. These are crucial for the explanation of the plateau-like features of $n_2(t)$ (see the text). The units are chosen such that $\hbar = 1 = e$.*

nation of these tendencies temporarily stops the depletion process on the central site leading to the plateaus. These time regimes are marked with A in Fig. 12. For times in which minima in the left lead's current coincide with an increase of $n_1(t)$, marked with B in Fig. 12, the depletion of the central site is even enhanced compared to the noninteracting case. The comparison to the curve for $U/\Gamma = 1$ but $V_{\rm b} = 0$ (dashed-dotted line in the upper panel) shows that a finite bias voltage is vital for the development of the plateaus. We note that only a few, mostly numerical, quantum many-body methods are available to study problems of this type.
The funRG approach to real-time dynamics was in addition used to study closed systems [35] as opposed to the open ones (due to the infinite reservoirs) considered here.

7 Summary

The present lecture had two main goals: (A) introducing the functional renormalization group approach to quantum many-body physics and (B) exemplifying how it is used as a versatile tool to study correlated quantum dots in and out of equilibrium. Over the last decade funRG was established as a modern tool to study quantum many-body systems with competing energy scales often resulting as a consequence of sizable correlations [10]. We introduced the method considering two toy models you have encountered before namely the classical and quantum

harmonic oscillator with quartic perturbation. By comparison to numerically exact results it was possible to show that the resummation inherent to the truncated funRG procedures leads to accurate results even beyond the limit of small perturbations. This holds despite the fact that the necessary truncations rendering the funRG an approximate method are guided by perturbation theory. Considering the single impurity Anderson model as well as the interacting resonant level model we discussed the prospects and limitations of the funRG approach to correlated quantum dots. Our particular focus was on transport through such systems either considering linear response or bias voltage driven nonequilibrium. An application of funRG to two-dimensional correlated materials is presented by C. Honerkamp in lecture C5.

References

[1] G.D. Mahan, *Many-particle Physics* (Plenum Press, New York, 1990)

[2] J.W. Negele and H. Orland, *Quantum Many-Particle Systems* (Addison-Wesley, Reading, 1988)

[3] H. Bruus and K. Flensberg, *Many-Body Quantum Theory in Condensed Matter Physics* (Oxford University Press, Oxford, 2004)

[4] J. Sólyom, Adv. Phys. **28**, 201 (1997)

[5] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 1999)

[6] R. Shankar, Rev. Mod. Phys. **66**, 129 (1994)

[7] A.C. Hewson, Adv. Phys. **43**, 543 (1994)

[8] F. Wegner, Annalen der Physik (Leipzig) **3**, 77 (1994)

[9] S. Kehrein, *The Flow Equation Approach to Many-Particle Systems* (Springer, Berlin, 2006)

[10] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, Rev. Mod. Phys. **84**, 299 (2012)

[11] T.A. Costi, Phys. Rev. B **64**, 241310(R) (2001)

[12] V. Meden and F. Marquardt, Phys. Rev. Lett. **96**, 146801 (2006)

[13] H.A. Nilsson, O. Karlström, M. Larsson, P. Caroff, J.N. Pedersen, L. Samuelson, A. Wacker, L.-E. Wernersson, and H.Q. Xu, Phys. Rev. Lett. **104**, 186804 (2010)

[14] A. Yacoby, M. Heiblum, D. Mahalu, and H. Shtrikman, Phys. Rev. Lett. **74**, 4047 (1995)

[15] C. Karrasch, T. Hecht, A. Weichselbaum, Y. Oreg, J. von Delft, and V. Meden, Phys. Rev. Lett. **98**, 186802 (2007)

[16] C. Karrasch, A. Oguri, and V. Meden, Phys. Rev. B **77**, 024517 (2008)

[17] A. Eichler, R. Deblock, M. Weiss, C. Karrasch, V. Meden, C. Schönenberger, and H. Bouchiat, Phys. Rev. B **79**, 161407(R) (2009)

[18] R. Hedden, V. Meden, Th. Pruschke, and K. Schönhammer, J. Phys.: Condens. Matter **16**, 5279 (2004)

[19] C. Karrasch, R. Hedden, R. Peters, Th. Pruschke, K. Schönhammer, and V. Meden J. Phys.: Condens. Matter **20**, 345205 (2008)

[20] C. Karrasch, V. Meden, and K. Schönhammer, Phys. Rev. B **82**, 125114 (2010)

[21] R. Gezzi, Th. Pruschke, and V. Meden, Phys. Rev. B **75**, 045324 (2007)

[22] S.G. Jakobs, M. Pletyukhov, and H. Schoeller, Phys. Rev. B **81**, 195109 (2010)

[23] J. Eckel, F. Heidrich-Meisner, S.G. Jakobs, M. Thorwart, M. Pletyukhov, and R. Egger, New J. Phys. **12**, 043042 (2010)

[24] M. Kinza, J. Ortloff, J. Bauer, and C. Honerkamp, Phys. Rev. B **87**, 035111 (2013)

[25] S. Streib, A. Isidori, and P. Kopietz, Phys. Rev. B **87**, 201107(R) (2013)

[26] Karrasch, Ph.D. thesis, RWTH Aachen University, Aachen, Germany, 2010 [http://d-nb.info/100962234X]

[27] D.M. Kennes, S.G. Jakobs, C. Karrasch, and V. Meden, Phys. Rev. B **85**, 085113 (2012)

[28] D. Bohr and P. Schmitteckert, Phys. Rev. B **75**, 241103 (2007)

[29] C. Karrasch, M. Pletyukhov, L. Borda, and V. Meden, Phys. Rev. B **81**, 125122 (2010)

[30] C. Karrasch, S. Andergassen, M. Pletyukhov, D. Schuricht, L. Borda, V. Meden, and H. Schoeller, Europhys. Lett. **90**, 30003 (2010)

[31] D.M. Kennes and V. Meden, Phys. Rev. B **87**, 075130 (2013)

[32] D.M. Kennes, D. Schuricht, and V. Meden, Europhys. Lett. **102**, 57003 (2013)

[33] H. Schoeller, Eur. Phys. J. Special Topics **168**, 179 (2009)

[34] D.M. Kennes and V. Meden, Phys. Rev. B **85**, 245101 (2012)

[35] D.M. Kennes and V. Meden, Phys. Rev. B **88**, 165131 (2013)

C 1 Magnetic phase transitions: from density functional theory to Monte Carlo simulations[1]

Phivos Mavropoulos

Institute for Advanced Simulation: Quantum Theory of Materials

Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Magnetic materials undergo in many cases a phase transition at some characteristic, critical temperature $T_{\rm c}$. For ferromagnets, $T_{\rm c}$ is the Curie temperature of the transition from the low-temperature ferromagnetic phase to high-temperature paramagnetic phase; for antiferromagnets, $T_{\rm c}$ is the Néel temperature where the antiferromagnetic order is lost in favour of the paramagnetic state again. Other magnetic phase transitions are known, e.g. between the antiferromagnetic and ferromagnetic state, or between non-collinear and paramagnetic states; the bottom line is, however, that if some type of magnetic order is present at low temperatures, then the temperature-induced fluctuations of the magnetization will kill it at high temperatures, unless of course the material melts before this point. The low-temperature order may be simple, as in a ferromagnet or antiferromagnet, or complex, as in the case of non-collinear magnetism, but it is always (by definition) characterized by an *order parameter* that is not always obvious and must be found by observation.[2]

Transitions between different types of order can occur as a function of the temperature; for example there is the notorious case of FeRh that has an antiferromagnetic ground state, turns to a ferromagnetic state at a temperature of about 370K, and has then a ferromagnetic-paramagnetic transition at $T_{\rm c} = 670$K. The mechanism underlying such transitions could be either solely electronic in nature (i.e., arising solely from fluctuations of the electron spins) or could be related to a structural phase transition or to external fields. They can be of first order, of second order, as the usual ferromagnetic-paramagnetic transition at the Curie point, or even higher orders. Their theoretical modelling and treatment can be a highly non-trivial task, with a program involving an understanding of the important degrees of freedom and the order parameter, the development of a statistical-mechanical model, the derivation of the interaction parameters for the model and the solution of the model equations.

For certain classes of problems the program has been developed and widely applied with rather good success. Such a class is the one of the materials showing a second order ferromagnetic-paramagnetic or antiferromagnetic-paramagnetic phase transition. Examples are the elemental ferromagnets Fe, Co, Ni or Gd, many transition metal alloys such as FeCo, or Heusler alloys of transition elements with non-magnetic elements, e.g. Co_2MnSi. In the present manuscript we will discuss their modelling, starting from the electronic structure calculations, the degrees of freedom relevant for magnetic order, the statistical-mechanical model based on the Heisenberg Hamiltonian and its extensions and its solution. Needless to say that the full program is not applicable to every ferromagnet, but it works for surprisingly many, taking into account the approximations involved.

Starting with an electronic structure method, and given the statistical-mechanical model, two main extensions are necessary from a computational point of view. The one is to extract the magnetic excitation energies from the electronic structure calculations. The second is to solve the model equations. Here we discuss both, emphasizing the utilization of the force theorem and Green-function based methods to reduce the numerical load of first task and the utilization of the Monte Carlo method for the second task.

[2] Although there is no prescription to find the order parameter in a system, one can infer the existence of a phase transition at some temperature by observing a divergence of the specific heat.

2 Energy scales, time scales and the Heisenberg model

We begin the discussion by motivating the use of a classical Heisenberg model for the statistical-mechanical description of magnetic systems. First we introduce the classical Heisenberg Hamiltonian that reads

$$H_{\mathrm{Heis}}[\{\boldsymbol{M}_i\}] = -\sum_{nn'} J_{nn'} \boldsymbol{M}_n \cdot \boldsymbol{M}_{n'} \tag{1}$$

where $J_{nn'}$ are interaction parameters and $\boldsymbol{M}_n$ is the magnetic moment at the atomic site n. It is assumed that the moments are only allowed to rotate and their modulus is constant. The negative sign is a usual convension leading to ferromagnetic interaction for $J_{nn'} > 0$ and antiferromagnetic for $J_{nn'} < 0$. The interactions are real and symmetric, $J_{nn'} = J_{n'n}$ while by definition $J_{nn} = 0$. The double sum implies double-counting.

Now we discuss under which physical assumptions Eq. (1) is valid. For the sake of simplicity we consider a ferromagnet at $T = 0$. If we wish to describe its magnetic behaviour up to the Curie temperature we should first estimate the degrees of freedom relevant for magnetism up to T_{c}, disregarding the rest. The magnetization is expected to be dominant in a certain region around the nuclei, within the so-called muffin-tin sphere where the wave-functions are most localized, while it should be small or negligible in the interstitial region. This motivates the concept of localized atomic magnetic moments in a material that are represented by $\boldsymbol{M}_n$ in Eq. (1). The next question is what type of magnetic fluctuations occur at low energies. We distinguish two types, the longitudinal fluctuations that are induced by suppressing the length of the atomic moments and the transverse fluctuations that are induced by tilting the directions of the atomic moments with respect to each other. Of the two, in the ferromagnets that we consider here, the transverse fluctuations are formed at a lower energy scale, rendering the tilting of the moments the low-lying type of excitation. Of course there are weak ferromagnets where the moments are small (e.g. $ZrZn_2$ with $0.12\mu_{\mathrm{B}}$) and "soft" and where the longitudinal and the transverse fluctuations are of the same scale, but we will not discuss their modelling here. Instead, we will discuss systems where the high localization of d or f states causes the electrons to be trapped at an atomic site long enough for the exchange interaction to form a strong, stable moment. This discussion clarifies the motivation for choosing only the rotational degrees of freedom in Eq. (1).

Next we motivate the choice of a classical model where $\boldsymbol{M}_n$ is the expectation value of the moment and not an operator. This choice is related to time-scale arguments. The precession of the atomic moments that has a time scale of the order of ps is normally much slower than the inter-site electron hopping with a time scale of the order of fs and than the intra-atomic electronic motions that determine the magnitude of the moments. In this respect, the magnetic moments appear as inert objects to the mobile electrons that move along the lattice. In the large time scale the local, longitudinal fluctuations of the moments have time to be averaged out, contributing to the scattered electrons by their expectation value only. This approach is called the *adiabatic approximation* in a similar sense to the Born-Oppenheimer approximation decoupling the electron motion from the lattice vibrations.

This lign of arguments has been supported and mathematically founded by e.g. You and Heine [1], Staunton et al. [2], Antropov et al. [3], Halilov et al. [4], or Gebauer and Baroni [5]. It has proven very fruitful, since it paves the way to calculate the exchange interactions $J_{nn'}$ from electronic structure calculations, as outlined in the next section, with results comparing rather well with experiment.

The major drawback of using a classical Heisenberg model is the low-temperature behaviour.

It is well known that the quantized nature of the magnon energies leads to a modest drop of the magnetization as a function of temperature following Bloch's law, $m(T) - m(0) \propto -T^{3/2}$ for $T \to 0$. This is theoretically established within the quantum Heisenberg model and experimentally found to a very good approximation. However, in the classical Heisenberg model the drop of $m(T)$ is linear, i.e., $m(T) - m(0) \propto -T$ for $T \to 0$, since the low excitations described by Eq. (1) are not quantized and can be formed for arbitrarily low moment angles. This difference in the behaviour is, as it turns out, not confined to very low temperatures, but continues up to perhaps $\frac{2}{3}T_\mathrm{c}$, after which $m(T)$ drops rather abruptly in the quantum model. Still, the fact that there is no obvious way to extract the interaction parameters for a quantum model from electronic structure calculations,[3] while density-functional methods make this possible for the classical model with reasonable success in predicting T_c makes the classical model popular and practically applicable. The successful prediction of T_c also suggests that the quantum nature of the magnetic excitations ceases to play a role at temperatures close to or above T_c.

3 From density functional theory to the Heisenberg model

We turn now to the way of calculation of the exchange parameters $J_{nn'}$ by means of density-functional calculations. For this we need the energy of the system in a number of excited states and in the ground state in order to fit the difference to Eq. (1). Since density-functional calculations provide only the ground state energy in principle, we resort to the adiabatic approximation that we discussed in the previous section. Assuming for simplicity of the discussion that we are examining a ferromagnet, the excited states that are relevant for the problem will be characterized by non-collinear atomic moments. In order to describe a non-collinear state within the adiabatic approximation we need to freeze the moments to form this state as if it were the ground state and use density-functional theory to calculate the energy. In order to freeze the moments, an additional *constraining field* $\boldsymbol{B}_n^\mathrm{constr}$ is applied at each atomic cell n and is adjusted in magnitude and direction so that the desired non-collinear state results. The direction of $\boldsymbol{B}_n^\mathrm{constr}$ is chosen to be perpendicular to the desired direction of $\boldsymbol{M}_n$ in order to avoid any energy contribution of the type $\boldsymbol{B}_n^\mathrm{constr} \cdot \boldsymbol{M}_n$. At the end, the density and spin density of the constrained state are calculated self-consistently using the usual Kohn-Sham procedure where the $\{\boldsymbol{B}_n^\mathrm{constr}\}$ are treated as additional external potentials. From this follows the energy of the constrained state, $E_\mathrm{CDFT}[\{\boldsymbol{M}_n\}]$, which in the adiabatic approximation is identified to the energy of the excited state. Thus the philosophy of the method is to substitute the ground state energy of an appropriately designed constrained system for the excited state energy of the original magnetic system and to accept the difference $E_\mathrm{CDFT}[\{\boldsymbol{M}_n\}] - E_\mathrm{DFT}$, where E_DFT is the ground-state energy of the unconstrained ferromagnet, as the excitation energy.

This philosophy is the basis of *constrained density functional theory* introduced by Dederichs et al. [7] that is mathematically founded on the concept of a constrained energy functional that, it the case that we are examining here, reads

$$F_\mathrm{CDFT}[\rho, \boldsymbol{m}; \{\boldsymbol{B}_n^\mathrm{constr}\}] = F_\mathrm{DFT}[\rho, \boldsymbol{m}] - \sum_n \int_{\Omega_n} d^3r\, \boldsymbol{B}_n^\mathrm{constr} \cdot [\boldsymbol{m}(\boldsymbol{r}) - \boldsymbol{M}_n] \qquad (2)$$

where F_DFT is the normal energy functional of the density ρ and spin-density $\boldsymbol{m}$ and Ω_n is the volume of the n^th atom in which the constraining field $\boldsymbol{B}_n^\mathrm{constr}$ is acting. From Eq. (2) one can

[3] If the classical parameters are used in the quantum Heisenberg model, then the T_c is strongly overestimated compared to experiment [6].

see that $\boldsymbol{B}_n^{\mathrm{constr}}$ acts formally as a Lagrance parameter; minimization of F_{CDFT} with respect to ρ and $\boldsymbol{m}$ leads to the solution $\int_{\Omega_n} d^3r\, \boldsymbol{m}(\boldsymbol{r}) = \boldsymbol{M}_n$ and the second term vanishes in the CDFT ground state. The minimization is done via the well-known Raleygh-Ritz method leading to Kohn-Sham equations with the additional potential term $V_{\mathrm{CDFT}} = -\sum_n \theta_n(\boldsymbol{r})\boldsymbol{B}_n^{\mathrm{constr}} \cdot \boldsymbol{\sigma}$, where $\theta_n(\boldsymbol{r})$ confines the action of $\boldsymbol{B}_n^{\mathrm{constr}}$ inside the n^{th} atomic sphere; in other words the usual magnetic exchange-correlation field $\boldsymbol{B}(\boldsymbol{r})$ is simply replaced by $\boldsymbol{B}(\boldsymbol{r}) + \sum_n \theta_n(\boldsymbol{r})\boldsymbol{B}_n^{\mathrm{constr}}$. It should be noted, however, that the input parameters in the calculations are the $\boldsymbol{B}_n^{\mathrm{constr}}$ and not the $\boldsymbol{M}_n$, therefore the $\boldsymbol{B}_n^{\mathrm{constr}}$ must be adjusted to obtain the desired $\boldsymbol{M}_n$.
Following this procedure one arrives at the aforementioned energy difference of the two states $E_{\mathrm{CDFT}}^{\mathrm{mag}}[\{\boldsymbol{M}_i\}] := E_{\mathrm{CDFT}}[\{\boldsymbol{M}_i\}] - E_{\mathrm{DFT}}$ that has to be compared to the Heisenberg model, Eq. (1). Examining if the form of $E_{\mathrm{CDFT}}^{\mathrm{mag}}$ for several non-collinear states fits Eq. (1) is a first test of the validity of the Heisenberg model. It frequently happens that Eq. (1) fails at high angles between nearest neighbours, in particular for d systems since the electron hopping can reduce the moment (in f systems the effect is much smaller because of the higher localization of the f wavefunctions). In this case one must either reconsider the model or further justify it. An argument for the use of the model even in the case of failure at high angles comes from Liechtenstein et al. [8] who argue that the magnetic state even at the Curie temperature is characterized by a non-vanishing correlation of neighbouring moment directions; in other words, some degree of ferromagnetic short range order is present at or above T_{c}, even if the long-range order is lost. This means that neighbouring moments will in general not reach too high angles with respect to each other, so that the Heisenberg model will still be a reasonable approximation to the energy $E_{\mathrm{CDFT}}^{\mathrm{mag}}$. This assumption should be tested in practice in each case. The argument, however, shows also that the parameters $J_{nn'}$ should not be calculated by comparing the ferromagnetic and antiferromagnetic state (which requires no constraining field and can be done within a collinear-magnetism calculation) but by comparing low-angle, non-collinear states, because the antiferromagnetic state could be outside the regime of validity of the Heisenberg model but is anyhow beyond the regime of possible configurations that the system is likely to reach even above T_{c}. From this perspective, going to the low-angle limit for the calculation of $E_{\mathrm{CDFT}}^{\mathrm{mag}}$, we may write

$$J_{nn'} = -\frac{\partial^2 H_{\mathrm{Heis}}[\{\boldsymbol{M}_i\}]}{\boldsymbol{M}_n \boldsymbol{M}_{n'}} = -\frac{\partial^2 E_{\mathrm{CDFT}}^{\mathrm{mag}}[\{\boldsymbol{M}_i\}]}{\boldsymbol{M}_n \boldsymbol{M}_{n'}} \tag{3}$$

Here the derivative with respect to $\boldsymbol{M}_n$ and $\boldsymbol{M}_{n'}$ is to be interpreted as a derivative with respect to the directions only, not the magnitude. The first equation follows directly from the Eq. (1). The second equation follows from the assumption that the constrained-density-functional energy landscape with respect to the moments is well approximated by the Heisenberg model at low angles. Thus the way is paved to fit the constrained-density-functional results to the Heisenberg model. Naturally the fitting does not necessarily have to be done by taking a derivative. Other methods, such as calculating $E_{\mathrm{CDFT}}^{\mathrm{mag}}[\boldsymbol{q}]$ for spin-spiral states of wave-vector $\boldsymbol{q}$ in reciprocal space and making a Fourier transform to obtain $J_{nn'}$ [4] have proven successful. One may also avoid the Fourier transform altogether and calculate $J_{nn'}$ as in Ref. [10] by a least-square fit to a number of $E_{\mathrm{CDFT}}^{\mathrm{mag}}[\boldsymbol{q}]$ at a reduced number of $\boldsymbol{q}$-points without necessity to cover the whole Brillouin zone, reducing computational time.[4]
The realization that the low-angle limit is physically correct leads to a further simplification. It is a good approximation at low angles to assume that the local moment is parallel to the

[4] In principle one needs a number of independent $\boldsymbol{q}$-states no greater than the number of independent $J_{nn'}$ that need to be calculated. Increasing the number of $\boldsymbol{q}$-states betters the quality of the fit.

local exchange plus constraining field, i.e., $\boldsymbol{M}_n \parallel \boldsymbol{B}_n + \boldsymbol{B}_n^{\mathrm{constr}}$ where $\boldsymbol{B}_n := \int_{\Omega_n} d^3r\, \boldsymbol{B}(\boldsymbol{r})$. This is because at low angles the "restoring force" of the moment to the equilibrium position is proportional to the angle, since the energy behaves quadratically around equilibrium, so that it can be neglected with respect to $\boldsymbol{B}^{\mathrm{constr}}$. Then the derivative in the last term of Eq. (3) may be replaced by

$$\frac{\partial^2 E^{\mathrm{mag}}_{\mathrm{CDFT}}}{\partial \boldsymbol{M}_n \partial \boldsymbol{M}_{n'}} = \frac{1}{M_n M_{n'}} \frac{\partial^2 E^{\mathrm{mag}}_{\mathrm{CDFT}}}{\partial \boldsymbol{e}_n \partial \boldsymbol{e}_{n'}}. \tag{4}$$

with $\boldsymbol{e}_n$ and $\boldsymbol{e}_{n'}$ unit vectors in the directions of $\boldsymbol{B}_n$ and $\boldsymbol{B}_{n'}$, respectively. The advantage of this approximation is that one avoids having to adjust the $\boldsymbol{B}_n^{\mathrm{constr}}$ to achieve the desired direction of $\boldsymbol{M}_n$ by a number of self-consistent calculations, thus reducing the computational time. The small discrepancy that is still present can be easily corrected for, as has been suggested by Bruno [14].

A further reduction of computational time comes from the use of the *magnetic force theorem*. The force theorem for density-functional calculations was introduced by Mackintosh and Andersen [15] as an approximation to total energy differences when displacing atoms in the crystal and its magnetic counterpart was introduced by Oswald et al. [16] for the calculation of exchange interactions in magnetic impurity pairs in nonmagnetic hosts. Loosely speaking, it states that the energy difference ΔE between two ground states that differ by an external parameter may be well approximated by calculating only one state, say state a, self-consistently, then modifying the potential so that it rigidly follows the change of the external parameter into the second state (b), calculating only the sum single-particle energies $E_{\mathrm{sp}}(b)$ without self-consistency, and finally subtracting the sum of single-particle energies of the first state; i.e., the theorem states that $\Delta E \approx E_{\mathrm{sp}}(b) - E_{\mathrm{sp}}(a)$ with $E_{\mathrm{sp}}(b)$ calculated non-self-consistently by rigidly modifying the potential of a. In terms of atom displacements the rigid displacement of the potential is represented by displacing the whole atomic sphere of the atom along with the nuclear position $\boldsymbol{r}_{\mathrm{N}}$ but keeping the form of the potential around the nucleus $V(\boldsymbol{r} - \boldsymbol{r}_{\mathrm{N}})$ constant. In terms of magnetic states the rigid displacement is a displacement only in spin space and amounts to calculating the ferromagnetic state (state a) and then flipping the $\boldsymbol{B}$-field in the atomic sphere of one of the atoms to obtain state b. The proof of the force theorem is based on the extremal properties of the energy functional and on the assumption that the density resulting from the rigid shift, $\rho_{\mathrm{rigid}}(b)$, is already a good approximation to the self-consistent density $\rho_{\mathrm{sc}}(b)$ making the error in the energy difference of order $(\rho_{\mathrm{rigid}}(b) - \rho_{\mathrm{sc}}(b))^2$ small. For magnetic systems, the assumptions of the force theorem are satisfied if $\boldsymbol{M}_n$ is parallel to $\boldsymbol{B}_n + \boldsymbol{B}_n^{\mathrm{constr}}$, which is the case if one calculates the ferromagnetic-antiferromagnetic energy difference, as in Ref. [16], or is approximately the case if the tilting angle from the ground state is small. The computational advantage due to the force theorem is that only one self-consistent calculation has to be performed, the one for the ground state, and the single-energy levels of all other states need to be calculated by their initial potential only without self-consistency.

Utilizing the magnetic force theorem, the small-angle approximation (4), and Eq. (3), we may derive a formula for the calculation of the interaction parameters $J_{nn'}$ that was given in an equivalent form by Liechtenstein et al. [8]. The strategy underlying this formula is called *method of infinitesimal rotations*. For the derivation we will need the concept of Green functions and a result from scattering theory that was presented in Chapter A5, Section 3.5 of the present volume. In the spirit of Eq. (3) we consider the two systems whose energy difference we need to calculate as (a) the ferromagnetic system and (b) the same system but with the moments at n and n' tilted by an infinitesimal amount by respective angles $\delta\alpha$ and $\delta\alpha'$ around given respective axes $\boldsymbol{u}_n$ and $\boldsymbol{u}_{n'}$ (which should be perpendicular to the ground-state magnetization). The rotation of

the moments is achieved, according to Eq. (4), by rotating the directions of the local $\boldsymbol{B}$-fields, $\boldsymbol{e}_n$ and $\boldsymbol{e}_{n'}$, by the same angles. The rotation of the potential $\boldsymbol{B}_n \cdot \boldsymbol{\sigma}$ is expressed as

$$\begin{aligned} \boldsymbol{B}_n \cdot \boldsymbol{\sigma} &\rightarrow e^{i\frac{\delta\alpha}{2}\boldsymbol{u}_n\cdot\boldsymbol{\sigma}}(\boldsymbol{B}_n \cdot \boldsymbol{\sigma})e^{-i\frac{\delta\alpha}{2}\boldsymbol{u}_n\cdot\boldsymbol{\sigma}} \\ &\approx \boldsymbol{B}_n \cdot \boldsymbol{\sigma} + i\,\frac{\delta\alpha}{4}\,[\boldsymbol{u}_n \cdot \boldsymbol{\sigma}, \boldsymbol{B}_n \cdot \boldsymbol{\sigma}] \\ &= \boldsymbol{B}_n \cdot \boldsymbol{\sigma} - \frac{\delta\alpha}{2}(\boldsymbol{u}_n \times \boldsymbol{B}_n) \cdot \boldsymbol{\sigma} \\ &= \boldsymbol{B}_n \cdot \boldsymbol{\sigma} - \frac{\delta\alpha}{2} B_n\, \boldsymbol{u}_{\perp n} \cdot \boldsymbol{\sigma}. \end{aligned} \tag{5}$$

where we expanded the exponential keeping only first order terms in $\delta\alpha$, $e^{\pm i\frac{\delta\alpha}{2}\boldsymbol{u}_n\cdot\boldsymbol{\sigma}} \approx 1 \pm i\frac{\delta\alpha}{2}\boldsymbol{u}_n\cdot\boldsymbol{\sigma}$, in the second step we used the property of the Pauli matrices, $(\boldsymbol{a}\cdot\boldsymbol{\sigma})(\boldsymbol{b}\cdot\boldsymbol{\sigma}) = \boldsymbol{a}\cdot\boldsymbol{b} + i(\boldsymbol{a}\times\boldsymbol{b})\cdot\boldsymbol{\sigma}$ for arbitrary vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, and in the third step we defined the unit vector $\boldsymbol{u}_{\perp n} = \boldsymbol{u}_n \times \boldsymbol{e}_n$ in the direction of $\boldsymbol{u}_n \times \boldsymbol{B}_n$ and implied $B_n(\boldsymbol{r}) = \boldsymbol{B}_n(\boldsymbol{r}) \cdot \boldsymbol{e}_n$ (to a good approximation, $\boldsymbol{B}_n$ is collinear within the atomic cell). The quantity $(\delta\alpha\, B_n\, \boldsymbol{u}_{\perp n})$ is just the constraining field $\boldsymbol{B}_n^{\mathrm{constr}}$, while $\delta\alpha\, \boldsymbol{u}_{\perp n} = \delta\boldsymbol{e}_n$. Repeating this for the field at n' we arrive at the result for the potential difference δV:

$$\begin{aligned} \delta V &= -\frac{\delta\alpha}{2} B_n\, \boldsymbol{u}_{\perp n} \cdot \boldsymbol{\sigma} - \frac{\delta\alpha'}{2} B_{n'}\, \boldsymbol{u}_{\perp n'} \cdot \boldsymbol{\sigma} \\ &=: \delta V_n + \delta V_{n'} \end{aligned} \tag{6}$$

with obvious abbreviations δV_n and $\delta V_{n'}$. This potential difference causes a change on the density of states $n(E)$ and on the integrated density of states $N(E) = \int^E dE'\, n(E')$ that are directly related to the sum of single particle energies. Taking into account a convolution with the Fermi function, we have:

$$\begin{aligned} E_{\mathrm{sp}} &= \int dE f(E) E\, n(E) \\ &= \Big[f(E)\, E\, N(E)\Big]_{-\infty}^{+\infty} - \int dE\, (f'(E)\, E + f(E))\, N(E) \\ &= E_{\mathrm{F}}\, N(E_{\mathrm{F}}) - \int^{E_{\mathrm{F}}} dE\, N(E) \qquad \text{for } T \to 0. \end{aligned} \tag{7}$$

where we performed an integration by parts, used the fact that $N(-\infty) = 0$ and $f(+\infty) = 0$, and that the derivative $-f'(E)$ approaches $\delta(E - E_{\mathrm{F}})$ as $T \to 0$. Setting the Fermi energy as a reference (which we can choose without loss of generality), we have $E_{\mathrm{sp}} = \int dE f(E)(E - E_{\mathrm{F}})n(E)$ and the above expression is simplified to

$$E_{\mathrm{sp}} = -\int^{E_{\mathrm{F}}} N(E)\, dE. \tag{8}$$

According to the force theorem, we must form the difference $\delta E_{\mathrm{sp}} = E_{\mathrm{sp}}(\delta\alpha, \delta\alpha') - E_{\mathrm{sp}}(0,0) = -\int^{E_{\mathrm{F}}} \delta N(E)\, dE$ (with obvious notation) in order to arrive at an expression for $J_{nn'}$. Using Eq. (3), Eq. (4) and the force theorem we have

$$J_{nn'} = \frac{1}{M_n M_{n'}} \frac{\int^{E_{\mathrm{F}}} \delta N(E)\, dE}{\delta\boldsymbol{e}_n\, \delta\boldsymbol{e}_{n'}}. \tag{9}$$

For the calculation of $\delta N(E)$ we use the result of scattering theory from Chapter A5, Section 3.5, which states that in the presence of a potential perturbation δV, the difference in the integrated density of states with respect to the unperturbed system is

$$\delta N(E) = \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\ln\mathcal{T}(E), \tag{10}$$

where

$$\mathcal{T}(E) = \delta V + \delta V\,\mathring{G}(E)\,\delta V + \delta V\,\mathring{G}(E)\,\delta V\,\mathring{G}(E)\,\delta V + \cdots \tag{11}$$

is the $\mathcal{T}$-matrix corresponding to the scattering potential δV with $\mathring{G}(E)$ the Green function of the unperturbed system. Since δV here is infinitesimal, its magnitude being controlled by the infinitesimal angles $\delta\alpha$ and $\delta\alpha'$, the expansion of Eq. (11) will be truncated to third order that is sufficient for the calculation of the interaction parameters, as we will see.
Inserting Eq. (11) into Eq. (10) we may use the rule $\mathrm{Tr}\log(AB) = \mathrm{Tr}\log A + \mathrm{Tr}\log B$ that is valid for matrices due to the trace even if A and B do not commute. We obtain

$$\delta N(E) = \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\ln\delta V + \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}\ln[1 + \mathring{G}(E)\,\delta V + \mathring{G}(E)\,\delta V\,\mathring{G}(E)\,\delta V + \cdots] \tag{12}$$

The first term on the rhs vanishes because of the imaginary part and because δV is hermitian making $\mathrm{Tr}\ln\delta V$ real. In the second part the logarithm may be expanded since δV is small (controlled by an infinitesimal real prefactor) yielding

$$\delta N(E) = \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}[\mathring{G}(E)\,\delta V] \;+\; \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}[\mathring{G}(E)\,\delta V\mathring{G}(E)\,\delta V] \;+\; \cdots \tag{13}$$

Inserting $\delta V = \delta V_n + \delta V_{n'}$ (Eq. (6)) we obtain the expansion

$$\begin{aligned}\delta N(E) \;=\; & \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}[\mathring{G}(E)\,\delta V_n + \mathring{G}(E)\,\delta V_{n'}] \\ & + \frac{1}{\pi}\mathrm{Im}\mathrm{Tr}[\mathring{G}(E)\,\delta V_n\,\mathring{G}(E)\,\delta V_n + \mathring{G}(E)\,\delta V_{n'}\,\mathring{G}(E)\,\delta V_{n'}] \\ & + \frac{2}{\pi}\mathrm{Im}\mathrm{Tr}[\mathring{G}(E)\,\delta V_n\,\mathring{G}(E)\,\delta V_{n'}] \\ & + \cdots\end{aligned} \tag{14}$$

This expression must be substituted into Eq. (9). On taking the derivatives with respect to the directions $\{\boldsymbol{e}_n\}$, or equivalently with respect to the parameters $\delta\alpha$ and $\delta\alpha'$ contained in δV_n and $\delta V_{n'}$, the various terms of Eq. (14) give different contributions either to $J_{nn'}$ or to other interaction parameters, on which we will briefly comment.

- The first term vanishes when taking the second derivative with respect to $\delta\alpha\delta\alpha'$. This term would be relevant only when taking the first derivative, giving a contribution to the energy dependence of a single moment, e.g. for the calculation of the anisotropy energy. However, since according to the assumptions of the force theorem one must start from a stationary state, i.e., at an energy minimum, in principle the first derivative should always vanish. Still, in some cases of non-collinear magnetism, especially in the presence of spin-orbit coupling, one applies the force theorem to the ferromagnetic state which is not the ground state, neglecting the possible error. Then the first term would have a contribution.

- The second term contributes only when taking the second derivative with respect to either $\delta\alpha^2$ or $\delta\alpha'^2$ and vanishes when taking the derivative with respect to $\delta\alpha\delta\alpha'$. This has no effect on the $J_{nn'}$ but contributes to the the anisotropy energy to lowest order, if one starts from the ground state. For a ferromagnet, the single-ion anisotropy will mainly include the exchange interaction summed over all neighbours; for a magnetic defect in a non-magnetic crystal the single-ion anisotropy will arise due to spin-orbit coupling.

- The third term is the only non-vanishing term when taking the derivative with respect to $\delta\alpha\delta\alpha'$, and is therefore the only one contributing to the $J_{nn'}$ according to Eq. (9).

- If the expansion had been carried out to higher orders in δV, then higher derivatives would be obtainable, containing also the Green function $\mathring{G}$ to higher order and corresponding to higher-order interactions, such as the bi-quadratic or the 4-spin interaction. These are, however, beyond the scope of the present chapter.

Returning to the calculation of $J_{nn'}$ we substitute the third term of Eq. (14) it into Eq. (9) and use the form of Eq. (6) as the perturbation δV. Remembering that $\delta\alpha\,\boldsymbol{u}_{\perp n} = \delta\boldsymbol{e}_n$ and similarly for n', we obtain the result

$$\begin{aligned} J_{nn'} &= \frac{1}{M_n M_{n'}}\frac{1}{2\pi}\mathrm{Im}\int^{E_\mathrm{F}} dE\,\mathrm{Tr}\big[\mathring{G}(E)\,B_n\boldsymbol{u}_{\perp n}\cdot\boldsymbol{\sigma}\,\mathring{G}(E)\,B_{n'}\boldsymbol{u}_{\perp n'}\cdot\boldsymbol{\sigma}\big] \\ &= \frac{1}{M_n M_{n'}}\frac{1}{2\pi}\int_{\Omega_n} d^3r\int_{\Omega'_n} d^3r'\,\mathrm{Im}\int^{E_\mathrm{F}} dE \\ &\quad\times\mathrm{Tr}_s\left[\mathring{G}(\boldsymbol{r}',\boldsymbol{r};E)\,B_n(\boldsymbol{r})(\boldsymbol{u}_{\perp n}\cdot\boldsymbol{\sigma})\,\mathring{G}(\boldsymbol{r},\boldsymbol{r}';E)\,B_{n'}(\boldsymbol{r}')(\boldsymbol{u}_{\perp n'}\cdot\boldsymbol{\sigma})\right]. \end{aligned} \tag{15}$$

In the last step we used a real-space representation whence the multiplication of matrices and the trace was converted into a space integrals in the volumes Ω_n and $\Omega_{n'}$ of cells n and n', leaving only a trace over spins, Tr_s.
From the assumptions of the Heisenberg model, Eq. (1), the value of $J_{nn'}$ does not depend on the direction of variation of the moments $\boldsymbol{M}_n$ and $\boldsymbol{M}_{n'}$. Due to this assumption it is sometimes said that the Heisenberg model is *isotropic*. Although this assumption is not always true in magnetic materials, and the resulting anisotropy is also found in electronic structure calculations, we assume for the moment that the assumption is valid, which is true if one can neglect the spin-orbit coupling or other causes for orbital polarization in the system. Under this assumption it does not matter which direction we choose for $\boldsymbol{u}_{\perp n}$ and $\boldsymbol{u}_{\perp n'}$, as long as they are perpendicular to $\boldsymbol{e}_n$. In the ground state of a ferromagnet or antiferromagnet that shows collinear magnetization we may also choose the direction of moments to be along the z axis and the directions of tilting of the moments equal to each other, $\boldsymbol{u}_{\perp n} = \boldsymbol{u}_{\perp n'}$, and along the x axis. Then we may set $\boldsymbol{u}_{\perp n'}\cdot\boldsymbol{\sigma} = \sigma_x$ in Eq. (15). In addition, the Green function of a collinear magnetic state with $\boldsymbol{B}(\boldsymbol{r}) \parallel z$ is diagonal in spin space with elements $\mathring{G}^{\uparrow}(E)$ and $\mathring{G}^{\downarrow}(E)$ and additionally the individual spin-matrix elements obey the symmetric property $\mathring{G}^s(\boldsymbol{r},\boldsymbol{r}';E) = \mathring{G}^s(\boldsymbol{r}',\boldsymbol{r};E)$. Then the action of $\sigma_x\mathring{G}(E)\sigma_x$ merely interchanges the elements $\mathring{G}^{\uparrow}(E)$ and $\mathring{G}^{\downarrow}(E)$. Under these conditions we arrive at the result

$$\begin{aligned} J_{nn'} &= \frac{1}{M_n M_{n'}}\int^{E_\mathrm{F}} dE\,\frac{1}{\pi}\,\mathrm{Im}\left[\int_{\Omega_n} d^3r\int_{\Omega'_n} d^3r'\mathring{G}^{\uparrow}(\boldsymbol{r}',\boldsymbol{r};E)\,B_n(\boldsymbol{r})\,\mathring{G}^{\downarrow}(\boldsymbol{r},\boldsymbol{r}';E)\,B_{n'}(\boldsymbol{r}')\right] \\ &=: \frac{1}{M_n M_{n'}}\int^{E_\mathrm{F}} dE\,j_{nn'}(E). \end{aligned} \tag{16}$$

This expression is simple to implement in any method that utilises the Green function. Based on it, one can delve into the microscopic background of the various exchange mechanisms.
Before discussing this matter, we will make a comment on Eq. (15). The integrand under $\int d^3r \int d^3r'$ is basically a special case of the magnetic susceptibility tensor $\chi_{ij}(\boldsymbol{r}, \boldsymbol{r}')$ that was derived in Chapter A5, Section 2 (Eq. 37) of the present volume. This realization suggests the following interpretation: a small constraining field at atom n' of the form $\delta \boldsymbol{B}^{\rm constr}_{n'}(\boldsymbol{r}') = \delta\alpha' B_{n'}(\boldsymbol{r}')\boldsymbol{u}_{\perp\, n'}$ (the particular form causes a moment rotation by $\delta\alpha'$) induces an additional magnetization at n of the form $\delta \boldsymbol{m}_n(\boldsymbol{r}) = \int_{\Omega_{n'}} d^3r' \chi(\boldsymbol{r}, \boldsymbol{r}') \delta \boldsymbol{B}(\boldsymbol{r}')$.[5] The latter interacts with a local constraining field at $\delta \boldsymbol{B}^{\rm constr}_{n}(\boldsymbol{r}) = \delta\alpha B_n(\boldsymbol{r})\boldsymbol{u}_{\perp\, n}$ via the energy term $\delta E_{nn'} = -\int_{\Omega_n} d^3r\, \delta \boldsymbol{m}_n(\boldsymbol{r}) \cdot \delta \boldsymbol{B}^{\rm constr}_{n}(\boldsymbol{r})$. Setting this interaction energy equal to[6] $-2J_{nn'}\, \delta \boldsymbol{M}_n \cdot \delta \boldsymbol{M}_{n'}$ gives us Eq. (15) for $J_{nn'}$.
Returning to the discussion of Eq. (16), the integrand $j_{nn'}(E)$ shows by how much the states at a certain energy contribute to the exchange interaction. The sign of $j_{nn'}(E)$ can change with energy, meaning that some states contribute to a ferromagnetic coupling and some contribute to an antiferromagnetic coupling. Thus it reveals competing interactions that in the end add up to the net value. Better intuition is provided by the integrals

$$J_{nn'}(E_1, E_2) = \int_{E_1}^{E_2} dE'\, j_{nn'}(E') \tag{17}$$

$$J_{nn'}(E) = \int_{-\infty}^{E} dE'\, j_{nn'}(E') = J_{nn'}(-\infty, E). \tag{18}$$

These energy-dependent exchange interactions show in a sense the value that the interaction would have if only some of the levels were filled (but without changing the band structure); the value of value of $J_{nn'}(E_{\rm F})$ is just the exchange parameter $J_{nn'}$. This line of thought gives a good starting point for analysis. For example, if a distinct group of states is recognized between E_1 and E_2, then $J_{nn'}(E_1, E_2)$ shows if and by how much these states contribute to the interaction. In another example, if $J_{nn'}(E)$ is close to a change of sign near $E_{\rm F}$, then we expect that small perturbations that can displace the bands or affect the chemical potential, e.g., external pressure or heavy doping, can potentially lead to a sign change of $J_{nn'}$. If the particular $J_{nn'}$ is known to be the most important for the ferromagnetic state, as is the case for the nearest-neighbour interaction in some ferromagnetic systems, then one can infer that small external perturbation can lead to a change of the magnetic ordering. Conversely, if the most important $J_{nn'}(E)$ is close to a local maximum near $E_{\rm F}$, then the ferromagnetic state is robust against external perturbations. We will see an example of $J_{nn'}(E)$ for bcc Fe in Section 5.
We close this section by noting that the exchange interaction, $J_{nn'}\, \boldsymbol{M}_n \cdot \boldsymbol{M}_{n'}$, is affected by the value of $J_{nn'}$ but also by the magnitude of the moments. The interaction between distant atoms is governed by both: $J_{nn'}$ gives the susceptibility $\chi_{nn'}$ of the electron liquid at atom n to a rotation of the moment at n' (if all other moments are held constant), while the magnitude of $\boldsymbol{M}_{n'}$ determines the amount of polarization induced through this susceptibility and the magnitude of $\boldsymbol{M}_n$ determines the interaction energy with the polarized liquid. In this sense, it is physically meaningful to consider the interaction as proportional to M_n and $M_{n'}$; in metallic systems it often shows an oscillating and decaying behaviour as a function of distance, analogous to the Rudermann-Kittel-Kasuya-Yoshida rule that is derived purely from the susceptibility due to the form of the Fermi surface [9]. On the other hand, for nearest neighbours the interaction

[5] We assume that the all moments other than n and n' are held constant.
[6] A prefactor 2 enters because we have both $J_{nn'}$ and $J_{n'n}$ present in Eq. (1).

is mediated by direct interaction of the atomic wavefunctions. Often, when mutually rotating the moments, chemical bonds that were present between the spin-up states or between the spin-down states of the neighbouring atoms are broken, while other bonds, between the spin-up states of one atom and the spin-down states of the other, are formed. The interaction strength is then governed by the energetics of bond formation and there is only little meaning in considering the magnitude of the moment, which could just as well be absorbed in the value of J.

4 The Monte Carlo method

We turn now to a different topic, namely the calculation of the temperature-dependent magnetization, susceptibility, and other thermodynamic quantities once the classical Hamiltonian (1) is given. The simplest way is to use mean-field theory [11]. The mean-field result for the critical temperature in a ferromagnet with one atom per unit cell is

$$k_{\mathrm{B}} T_{\mathrm{c}}^{\mathrm{mf}} = \frac{2}{3} \sum_{n'} J_{nn'}. \tag{19}$$

The mean field method has the benefit of being numerically inexpensive and giving usually correct trends, but also has severe shortcomings. It always overestimates the Curie temperature to the point that it gives non-zero T_{c} even in one-dimensional systems. While in three-dimensional close-packed systems it performs reasonably well, especially if the interactions are long-ranged, it fails for instance in diluted magnetic alloys with short-range interactions. Additionally, it provides no information on the fluctuations.

Alternatively there is the random phase approximation originally developed by Tyablikov [12] and Callen [13] that approaches much better the correct solution of the Heisenberg model but is not able to handle systems with a non-collinear ground state. It should be noted that the random phase approximation is a method for solving the quantum Heisenberg model $H = -\sum J_{nn'} \hat{\boldsymbol{s}}_n \cdot \hat{\boldsymbol{s}}_{n'}$, where $\{\hat{\boldsymbol{s}}_n\}$ are spin operators of a given spin quantum number s, however it can be used in for the classical model by taking the limit of large spin quantum number, $S \to \infty$ and scaling down the value of $J_{nn'}$ accordingly, so that the product $J_{nn'} s(s+1)$ remains constant.

A method to solve the Heisenberg model to arbitrary precision, albeit with increasing numerical cost, is the Monte Carlo method, which we will introduce in this section. It is a stochastic method based on the use of series of random numbers for the generation of Markov chains, i.e., a succession of states in a system where a state i is followed by another state $i+1$ based on a rule that depends only on state i and to a given probability distribution. In this way, the space of states of the system (also called *phase space*) is sampled step-by-step and at the end one may average the desired quantities over a large number of steps in order to obtain thermodynamical information. The name Monte Carlo was coined by Metropolis et al. [17] who were among the first to apply this method.

The Monte Carlo method has applications on a wide range of problems in statistical physics and even quantum many-body problems. Here we will give the general principles and discuss some applications in magnetism. A comprehensive introduction can be found in the books of Newman and Barkema [18] or Landau and Binder [19].

4.1 Preliminaries on statistically averaged physical quantities

In this section we remind the reader of a few basic formulas of statistical physics that are useful in the subsequent sections. We are interested in the statistical properties of a system at thermal equilibrium at temperature T. For the purpose of the present lecture notes the system is a magnetic material described by the classical Heisenberg Hamiltonian, Eq. (1). The probability that the system is at the $i^{\rm th}$ state with energy E_i is given by the Gibbs-Boltzmann distribution

$$p_i = \frac{1}{\mathcal{Z}}\, e^{-\frac{E_i}{k_{\rm B}T}} = \frac{1}{\mathcal{Z}}\, e^{-\beta\, E_i} \tag{20}$$

where we introduced $\beta = (k_{\rm B}T)^{-1}$ for later convenience. $\mathcal{Z}$ is the *partition function*

$$\mathcal{Z} = \sum_j e^{-\beta\, E_j} \tag{21}$$

where the sum, extending over all states j, should be re-interpreted as an integral in the case of a continuous phase space.[7] The quantity $F = \log \mathcal{Z}$ is called free energy. $\mathcal{Z}$ is evidently a function of all external parameters that define the dynamics of the system, e.g. the volume, the external fields, etc., since all energy levels E_j depend on these parameters. In the Gibbs-Boltzmann distribution (20) the partition function plays merely the role of a normalization constant but in reality it serves as a generating function for a number of thermodynamic quantities. For example, the average energy of the system, $\langle E\rangle := \sum_j E_j\, p_j$, is written as

$$\langle E\rangle = \frac{1}{\mathcal{Z}} \sum_j E_j\, e^{-\beta\, E_j} \tag{22}$$

$$= -\frac{1}{\mathcal{Z}}\frac{\partial \mathcal{Z}}{\partial \beta} = -\frac{\partial \log \mathcal{Z}}{\partial \beta} \tag{23}$$

while the average of the squared energy fluctuation amplitude is

$$\left\langle \left(E - \langle E\rangle \right)^2 \right\rangle = \left\langle E^2\right\rangle - \langle E\rangle^2 \tag{24}$$

$$= \frac{1}{\mathcal{Z}} \sum_j E_j^2\, e^{-\beta\, E_j} - \left(\frac{1}{\mathcal{Z}}\frac{\partial \mathcal{Z}}{\partial \beta}\right)^2$$

$$= \frac{1}{\mathcal{Z}}\frac{\partial^2 \mathcal{Z}}{\partial \beta^2} - \left(\frac{1}{\mathcal{Z}}\frac{\partial \mathcal{Z}}{\partial \beta}\right)^2 = \frac{\partial^2 \log \mathcal{Z}}{\partial \beta^2}. \tag{25}$$

On the rhs we recognize the derivative $-\partial \langle E\rangle\, \partial\beta$ according to Eq. (23). But since $\beta = (k_{\rm B}T)^{-1}$, this term is directly related to the specific heat $C_T = \partial \langle E\rangle / \partial T$. Thus we have the result

$$C_T := \frac{\partial \langle E\rangle}{\partial T} = k_{\rm B}\, \beta^2 \, \left(\left\langle E^2\right\rangle - \langle E\rangle^2 \right). \tag{26}$$

Observable quantities that couple linearly to external fields can be averaged following the same procedure. For instance, since the magnetization $\boldsymbol{m}(\boldsymbol{r})$ couples to an external magnetic field

[7] For example, in the case of the Heisenberg model (1), we may postulate $\sum_j \to \prod_n \int d\omega_n$, where $d\omega_n$ is the solid angle element of the $n^{\rm th}$ moment. If the moments are of different length we may postulate $\sum_j \to \prod_n M_n^2 \int d\omega_n$.

$\boldsymbol{B}_{\rm ext}$ by the adding the term $-\boldsymbol{B}^{\rm ext}\cdot\boldsymbol{M}^{\rm tot} := -\int d^3r \boldsymbol{B}^{\rm ext}\cdot\boldsymbol{m}(\boldsymbol{r})$ to the Hamiltonian, the energy levels are shifted as $E_i \to E_i - \boldsymbol{B}^{\rm ext}\cdot\boldsymbol{M}_i^{\rm tot}$, where $\boldsymbol{M}_i^{\rm tot}$ is the total magnetic moment of the $i^{\rm th}$ state. Let us consider for simplicity the component of $\boldsymbol{M}_i^{\rm tot}$ is in the direction of $\boldsymbol{B}^{\rm ext}$ and call this $M_{\parallel i}^{\rm tot}$(which is the case for the Heisenberg model (1) and for magnets without magnetic anisotropy). Then the average magnetic moment along $\boldsymbol{B}^{\rm ext}$ can be found by

$$\begin{aligned}\left\langle M^{\rm tot}\right\rangle &= \frac{1}{\mathcal{Z}(B^{\rm ext})}\sum_j M_{\parallel j}^{\rm tot} e^{-\beta\,(E_j - B^{\rm ext}M_{\parallel j}^{\rm tot})} \\ &= \frac{1}{\beta}\frac{\partial \log \mathcal{Z}(B^{\rm ext})}{\partial B^{\rm ext}} \end{aligned} \tag{27}$$

and the squared fluctuation amplitude by

$$\left\langle \left(M_\parallel^{\rm tot} - \left\langle M_\parallel^{\rm tot}\right\rangle\right)^2\right\rangle = \left\langle (M_\parallel^{\rm tot})^2\right\rangle - \left\langle M_\parallel^{\rm tot}\right\rangle^2 \tag{28}$$

$$\begin{aligned} &= \frac{1}{\mathcal{Z}(B^{\rm ext})}\sum_j (M_{\parallel j}^{\rm tot})^2\, e^{-\beta\,(E_j - B^{\rm ext}M^{\rm tot}\parallel)} - \left(\frac{1}{\beta}\frac{\partial \log \mathcal{Z}(B^{\rm ext})}{\partial B^{\rm ext}}\right)^2 \\ &= \frac{1}{\beta^2}\frac{\partial^2 \log \mathcal{Z}(B^{\rm ext})}{\partial (B^{\rm ext})^2} \end{aligned} \tag{29}$$

Again, on the rhs we recognize the presence of the derivative $\partial\left\langle M_\parallel^{\rm tot}\right\rangle/\partial B^{\rm ext}$ by virtue of Eq. (27). But this is the just the longitudinal magnetic susceptibility X_T. Thus we obtain the result, known as *linear response theorem*,

$$X_T := \frac{\partial\left\langle M_\parallel^{\rm tot}\right\rangle}{\partial B^{\rm ext}} = \beta\,\left[\left\langle\,(M_\parallel^{\rm tot})^2\,\right\rangle - \left\langle M_\parallel^{\rm tot}\right\rangle^2\right]. \tag{30}$$

Since the derivation of this expression did not explicitly depend on the fact that $M_\parallel^{\rm tot}$ is the magnetic moment and $B^{\rm ext}$ an external field, but only on the fact that they are linearly coupled in the Hamiltonian, the same expression is valid for the response function of any quantity to an external perturbation as long as the two are coupled linearly.

By virtue of Eq. (30) the susceptibility can be calculated via the fluctuation amplitude without resorting to numerical differentiation that requires an accurate calculation of $M_\parallel^{\rm tot}$ at a number of values of $B^{\rm ext}$. Since this expression is still valid at $B^{\rm ext}\to 0$, the fluctuation amplitude in the absence of a magnetic field suffices to find the zero-field susceptibility. Similarly, Eq. (26) allows the calculation of the specific heat without resorting to numerical differentiation. This leads to a significant reduction of computational time, especially in the Monte Carlo technique, since the numerical differentiation may require very precisely calculated values of the differentiated function, while the precision in the Monte Carlo method increases only logarithmically with computational time (number of sampling steps).

Other quantities whose calculation is simplified by this trick are the entropy $S = k_{\rm B}\log\Gamma$ (with Γ the volume of phase space compatible with a certain thermodynamical state) and the free energy $F = \langle E\rangle - TS$. In particular it is known from statistical physics that the entropy is related to the specific heat via the relation $C_T = T\partial S/\partial T = -\beta\partial S/\partial\beta$. On the other hand, C_T is most conveniently calculated by means of the energy fluctuation, Eq. (26). This suggests that the entropy is found by means of an integration,

$$S(T) = S(T_0) + \int_{T_0}^{T}\frac{C_T(T')}{T'}\,dT', \tag{31}$$

where it is assumed that S is known at some initial temperature T_0. In systems satisfying the third law of thermodynamics we have by assumption $S(0) = 0$, thus we may set $T_0 = 0$. This is not, however, the case in the Heisenberg model (1) that has a degenerate ground state with a continuous symmetry (all spins may be rotated by the same angle without changing the energy), thus $S(0) = -\infty$. The same problem occurs for all models that show a continuous symmetry in the ground state. Then one may choose $T_0 = \infty$ (in practice very large) using some approximation for $S(\infty)$ and integrate downwards to obtain $S(T)$. Once $S(T)$ is found, the free energy is calculated using $F = \langle E \rangle - TS$.

4.2 Basic principles of the Monte Carlo method

The Monte Carlo method is based on a sampling of the phase space by means of a stochastic process, i.e., a process where the next state is always chosen with a certain probability among many possible states. Each state found j during the process contributes to the quantity that is being averaged, e.g. the energy or magnetization. In the ideal case the rule of the stochastic process, i.e., the rule for constructing the transition probability w_{fi} from the initial state i to the final state f, should be constructed in such a way that the number of necessary steps in order to obtain a converged averaging is kept to a minimum. Thus for every different type of problem there might be a different optimal rule. Unfortunately there is no systematic way of optimizing the stochastic rule, so the optimization relies solely on ingenuity and intuition.
In the stochastic rule, w_{fi} depends only on the initial and final state, but not explicitly on any of the states that were visited prior to the initial. A stochastic rule is formally correct if it obeys two principles:

1. It should be ergodic, i.e., all states of the phase space should in principle be accessible.
2. It should obey the principle of *detailed balance*, i.e.,

$$w_{i \leftarrow j}\, p_j = w_{j \leftarrow i}\, p_i \tag{32}$$

   where p_i and p_j are the probabilities of occurrence of states i and j, respectively, in the statistical distribution. The arrows in the subscripts stress the initial and final states. Without the detailed balance condition it may be that the stochastic process overestimates or underestimates the probability of occurrence of some states resulting in wrong statistics.

In Section 4.3 we will discuss two stochastic rules, the Metropolis algorithm and the heat-bath algorithm.

Importance sampling and statistical averages

Having decided on some stochastic rule, we create a random walk in phase space by a sequence of stochastic processes. During this walk we sample the phase space, meaning that we calculate the quantities that we wish to average (e.g. energy or magnetic moment) at each visited state. Since the random walk will never have the time to visit all of phase space (or even to create an approximately dense trajectory in the full phase space), the stochastic rule should be optimized in such a way that the physically most important part of phase space is sampled. Therefore the name *importance sampling* has been coined for this procedure.
In the concrete example of a ferromagnet, we model the system by a supercell containing a large number of elementary cells and obeying periodic boundary conditions to avoid spurious surface

effects. The magnetization is then defined as the magnetic moment per atom, $m = M^{\rm tot}/N_{\rm at}$, where $M^{\rm tot}$ is the calculated total moment of the supercell (the summed moments of all atoms) and $N_{\rm at}$ the number of atoms in the supercell. We introduce an integer parameter t that is named *simulation time* and that basically counts the steps of the stochastic process.[8] At each step we define the magnetization $m(t)$; the full state at each t is sometimes called a *snapshot*. The thermal average of the magnetization is then found by averaging over a large simulation time interval $t_1 \leq t \leq t_2$. We simply have

$$\langle m \rangle = \frac{1}{t_2 - t_1 + 1} \sum_{t=1_1}^{t_2} m(t), \quad m(t) = \frac{M^{\rm tot}(t)}{N_{\rm at}}. \tag{33}$$

All observables are averaged in the same way which also provides derived quantities. For example, the susceptibility per atom, $\chi_T = X_T/N_{\rm at}$ is given, using Eq. (30), by the squared fluctuation amplitude of the magnetization:

$$\chi_T = \beta N_{\rm at} \left(\langle m^2 \rangle - \langle m \rangle^2 \right) \tag{34}$$

with

$$\langle m^2 \rangle = \frac{1}{t_2 - t_1 + 1} \sum_{t=1_1}^{t_2} m(t)^2. \tag{35}$$

Correlation functions

Other quantities that are of interest are correlation functions, $\langle AB \rangle$ showing how a quantity A is dependent upon a quantity B. For example, it can be of interest to see how much two magnetic moments $\boldsymbol{M}_n$ and $\boldsymbol{M}_{n'}$ are correlated in a ferromagnet, as a function of the distance between sites n and n'. Then we may build the average $G_{nn'}(T) = \langle \boldsymbol{M}_{n'} \cdot \boldsymbol{M}_{n'} \rangle$ that defines the *correlation function*, or the average $G^{(\rm c)}_{n\mu;n'\nu}(T) = \langle (M_{n;\mu} - \langle M_{n;\mu} \rangle)(M_{n';\nu} - \langle M_{n';\nu} \rangle) \rangle$ that defines the *connected correlation function* and gives the correlation between the moment fluctuations at n in the direction μ and at n' in the direction μ ($\mu, \nu \in x, y, z$), rather than between the moments themselves. Applying the same procedure that led us to the linear response theorem, Eq. (30), we find that $\chi_{n\mu;n'\nu}(T) = \beta G^{(\rm c)}_{n\mu;n'\nu}(T)$ is a susceptibility tensor describing the change in the moment $\boldsymbol{M}_n$ if an infinitesimal external field is applied at $\boldsymbol{M}_{n'}$. Intuitively this makes sense, since the fluctuations of the moments can affect each other if the application of a field at one moment affects the direction of the other moment.

Autocorrelation function

We proceed by introducing the concept of the autocorrelation function of a calculated quantity. As an example we use again the magnetization in the Heisenberg model of a ferromagnet. The question is how long the time interval Δt should be in Eq. (33) to obtain reliable results. Except the obvious fact that the fluctuations will decrease with increasing Δt, there is the additional issue that the system may stay in the vicinity of the initial state for a long time so that the phase space is not properly scanned. It can be that this is because there is an energy barrier between the initial state and other parts of the phase space, and the stochastic process needs time to overcome the barrier between two local minima. A second reason could be that there

[8] The simulation time t is not necessarily related to physical time, but certainly related to computational time.

is no barrier but the stochastic rule is just poorly designed for this case, keeping the system close a phase-space point for too long. For example, if the ferromagnet is close to $T = 0$ and the stochastic rule is to pick a moment at random and rotate it in a random direction, using the Metropolis criterion for acceptance (which is discussed in Section 4.3), then the resulting state will have a too high energy and will be rejected in most cases. Thus the system will remain long in the initial state, even though there is no barrier to overcome. In this sense, the state of the system remains correlated to states at earlier times.
The autocorrelation function quantifies this correlation. For the example of the magnetization it is defined as

$$\begin{aligned} c(t) &= \langle\, m(t')\, m(t'+t)\,\rangle - \langle m(t')\rangle\ \langle m(t'+t)\rangle_{t'} \\ &= \frac{1}{\Delta t}\sum_{t'=t_1}^{t_2} m(t')m(t'+t) - \left(\frac{1}{\Delta t}\sum_{t'=t_1}^{t_2} m(t')\right)\left(\frac{1}{\Delta t}\sum_{t'=t_1}^{t_2} m(t'+t)\right). \end{aligned} \qquad (36)$$

where the subscript $\langle\cdots\rangle_{t'}$ indicates an average over t' and $\Delta t = t_2 - t_1 + 1$. Effectively, $c(t)$ measures not the correlation between the time-displaced magnetization, but between the fluctuations of m around the equilibrium at different times displaced by t. To see this, we can assume that at equilibrium $\langle m(t')\rangle_{t'} = \langle m(t'+t)\rangle_{t'} = \langle m\rangle$ so that from the definition we obtain the relation $c(t) = \langle\, [m(t') - \langle m\rangle][m(t'+t) - \langle m\rangle]\,\rangle_{t'}$.
The autocorrelation function falls off approximately as $c(t) \propto e^{-t/\tau}$, where τ is called the autocorrelation time. After approximately 2τ, $c(t)$ reaches zero and fluctuates at small positive and negative values. At this point we can say that the state is not correlated to the initial state. Thus in calculating thermally averaged quantities we must sample for a time interval $\Delta t \gg \tau$.

Initial relaxation time

When starting a Monte Carlo simulation the very first state is rarely a good representative of thermal equilibrium. For instance, in order to study the state of a ferromagnet close to T_c, neither a ferromagnetic state nor a paramagnetic state with fully randomized directions of the moments can be regarded as a good approximation. The magnetic state at T_c comprises regions that appear nearly ferromagnetic, but with their magnetization pointing in different directions with respect to each other, as well as regions that appear paramagnetic, encompassing fluctuations at all length scales from the inter-atomic nearest neighbour distance atom to the full crystal (or simulation supercell in calculations). Such a state is next to impossible to design as an "initial guess," so it must be produced by simulation.
Clearly then, starting from a given state that does not reflect the equilibrium properties at T, one must let the simulation run for a number of steps before starting to sample, so that the system is able to reach thermal equilibrium. This necessary number of steps is called *initial relaxation time*. It is not the same as the autocorrelation time, since the latter refers to a starting point that corresponds already to equilibrium.
One way to estimate the initial relaxation time is to monitor the averaged quantities until they are stabilized around some mean value and then start the sampling. Using a trick, it is even possible not to waste any more Monte Carlo steps than necessary. In particular we observe (taking the magnetization as an example)

$$\frac{1}{t_2 - t_1 + 1}\sum_{t=t_1}^{t_2} m(t) = \frac{t_2}{t_2 - t_1 + 1}\frac{1}{t_2}\sum_{t=1}^{t_2} m(t) - \frac{t_1}{t_2 - t_1 + 1}\frac{1}{t_1}\sum_{t=1}^{t_1} m(t) \qquad (37)$$

or

$$\langle m \rangle_{(t_1,t_2)} = \frac{t_2}{t_2 - t_1 + 1} \langle m \rangle_{(1,t_2)} - \frac{t_1}{t_2 - t_1 + 1} \langle m \rangle_{(1,t_1)} \tag{38}$$

where we denoted by $\langle m \rangle_{(t,t')}$ the mean value of m averaged between t and t'. Evidently, if the values of $\langle m \rangle_{(1,t_1)}$ (and the other averaged quantities) are stored every few steps between 1 and t_2, we can use Eq. (38) to find at which value of t_1 the magnetization does not change any more and accept this as the initial relaxation time. Simultaneously we obtain the correct average. With this trick we avoid the need to repeat the calculation or to store the full state at every few steps.

Random number generation

Since each step of the stochastic process involves a probability distribution, a numerical implementation of the method requires a random number generator. This is the heart of any Monte Carlo algorithm and the correctness of the results depends critically on its quality. Usually a numerical simulation of a random number sequence is employed giving a uniform probability distribution in the semi-open interval $[0, 1)$. It is created by an algorithm and is fully reproducible and also periodic, but it appears to have the characteristics of "truly random" numbers, hence it bears the name *pseudo-random sequence*. In particular it should have a period long enough so as not to repeat itself during the stochastic process and it should show no correlation, meaning that the probability of obtaining a value should not detectably depend on the values previously obtained. Needless to say that the algorithm should also appear ergodic, i.e., densely cover the full interval $[0, 1)$ limited only by its period and by the truncation of decimal points in the computer.
In many applications there is the need for a non-uniform probability distribution following a function $p(x) \geq 0$ with $\int_a^b p(x)\,dx = 1$, where a can be $-\infty$ and b can be ∞. There is a method for obtaining random numbers x according to $p(x)$, if random numbers r are provided in the uniform distribution in $[0, 1)$. The method requires knowledge of the integral $P(x) = \int_a^x p(x')\,dx'$ and its inverse function, P^{-1}. Obviously, P is monotonously increasing and $P(a) = 0$, $P(b) = 1$; thus, P^{-1} is a monotonous function defined in the interval $[0, 1]$ and $P^{-1}(0) = a$, $P^{-1}(1) = b$. Then, if a random variable r obeys the uniform distribution in the interval $[0, 1)$, the random variable $P^{-1}(r)$ obeys the distribution p in $[a, b)$. This is proven e.g. in Refs. [18, 19].

4.3 Metropolis, heat-bath, and cluster-flipping algorithms

Metropolis algorithm

We now introduce one of the first and simplest stochastic rules for Monte Carlo calculations, the Metropolis algorithm [17]. For the Heisenberg model, it has as follows:

1. Starting from an initial state i, choose a moment at random.
2. Choose a candidate new state f by rotating the moment to a new direction at random.
3. Calculate the energy difference between the new and old state, $E_f - E_i$.
4. If $E_f - E_i < 0$ accept the candidate state as the next state. Otherwise compare a random number $r \in [0, 1)$ to $\exp\left[-\beta\left(E_f - E_i\right)\right]$; if $r < \exp\left[-\beta\left(E_f - E_i\right)\right]$ accept the state f

otherwise reject it and keep i as the next state. Perform a sampling at the resulting state (either f or i).

5. Using the resulting state (either i or f), return to step 1.

Basically, steps 3 and 4 are called *Metropolis criterion* and are the essence of the metropolis algorithm. Steps 1 and 2 have to do with the specific system (i.e., interacting magnetic moments), the way of partitioning phase space (one atomic moment at a time) and the principle of choosing the candidate new state f (rotate the moment to a randomly chosen direction).
The Metropolis algorithm satisfies the basic criteria for a stochastic rule. It is obviously ergodic, since all moments have a probability of being rotated in any direction, thus all final states have a probability of being accepted. It also satisfies the principle of detailed balance by construction of the Metropolis criterion. To see this we may write the detailed balance principle (32) as

$$\frac{w_{j\leftarrow i}}{w_{i\leftarrow j}} = e^{-\beta\,(E_j - E_i)}. \tag{39}$$

But this is exactly what is done by the Metropolis criterion: going lower in energy has probability 1 while going higher in energy has probability $e^{-\beta\,|\Delta E|}$.
The Metropolis algorithm is the workhorse of the Monte Carlo method. It is simple to implement and in many cases very efficient. However, one can make adjustments to improve it for particular problems by improving the *acceptance ratio*, which is the ratio of accepted states in step 4 to the total number of trials. Intuitively we expect that the optimal acceptance ratio is $\frac{1}{2}$. If it is much lower, the computational time is waisted in rejecting states that were probably irrelevant for the problem anyway; if it is too close to 1, then it means that the stochastic process could be failing to sample an important part of the phase space, being trapped at some local minimum of the free energy (although at very high temperatures the ratio must approach 1). A typical example for the former scenario is the case of a ferromagnet at low temperatures that we discussed when introducing the autocorrelation function: many new random moment directions increase the energy too much, posing the moment at a high angle to the neighbours, and are rejected. One simple but efficient way to circumvent this problem is to restrict the candidate new moment direction to a cone of fixed opening angle θ_0 around the previous direction. As θ_0 becomes smaller, the energy of the candidate state will be closer to the one of previous state, increasing the acceptance ratio. The optimal θ_0 is temperature-dependent; at a certain temperature we may adjust θ_0 by a few trials to optimize the acceptance ratio and then fix it and continue the simulation. A reasonable guess of θ_0 for an acceptance ratio of $\frac{1}{2}$ in the case of the Heisenberg model (1) at low temperatures is motivated by the demand that a rotated moment should give $e^{-\beta|\Delta E|} = \frac{1}{2}$, i.e., $\exp[(2\beta \sum_{n'} J_{nn'} M_n M_{n'})(1 - \cos\theta_0)] = \frac{1}{2}$,[9] or equivalently $\cos\theta_0 = 1 - [2\beta \sum_{n'} J_{nn'} M_n M_{n'}]^{-1} \log 2$.

Heat-bath algorithm

Next we discuss a stochastic rule of different philosophy, the *heat-bath algorithm*. It was first introduced for the Ising model but extended also to the Heisenberg model by Miyatake et al. [20]. The idea is to consider, at each step, a magnetic moment of the crystal $\boldsymbol{M}_n$ as if it were in contact with a heat bath of temperature T and under the action of a local Hamiltonian $H_n^{\rm loc} = -2\boldsymbol{M}_n \cdot \boldsymbol{B}_n^{\rm eff} := -2\boldsymbol{M}_n \cdot (\sum_{n'} J_{nn'} \boldsymbol{M}_{n'})$, where we have defined an effective field

[9] The factor 2 in the exponential originates from the double counting in Eq. (1).

acting on $\boldsymbol{M}_n$ as the sum of the interaction to all neighbours of the Heisenberg Hamiltonian (1). Being in this environment, the moment $\boldsymbol{M}_n$ will fluctuate, attaining various directions according to the Gibbs-Boltzmann distribution, $p(\boldsymbol{M}_n) = \exp\left[-\beta\left(-2\boldsymbol{M}_n \cdot \boldsymbol{B}_n^{\text{eff}}\right)\right]/\mathcal{Z}$, where the partition function comprises an integral over the solid angle: $\mathcal{Z} = \int d\omega_n \exp\left[-\beta\left(-2\boldsymbol{M}_n \cdot B_n^{\text{eff}}\right)\right]$. Given this form of $p(\boldsymbol{M}_n)$, the next step is to choose a new direction as a set of two random variables, the angles θ and ϕ, that obey this particular probability distribution. This new direction is always accepted; i.e., the heat-bath method has no rejection.

Using a local coordinate system with the z axis along $\boldsymbol{B}_n^{\text{eff}}$, the angle ϕ trivially obeys the uniform distribution in the interval $[0, 2\pi)$, since the energy is rotationally invariant around $\boldsymbol{B}_n^{\text{eff}}$. For the angle θ the situation is more complicated and one must resort to the transformation from the uniform distribution to the new distribution as was outlined at the end of Section 4.2. The resulting integral and function inversion was performed analytically by Miyatake et al. [20] yielding an angle θ_n that should be chosen according to the rule

$$\cos\theta_n = \frac{1}{\beta E_0} \log\left[(1-r)\exp(\beta E_0) + r\exp(-\beta E_0)\right] \tag{40}$$

where r is a random variable obeying the uniform distribution in $[0, 1)$ and $E_0 = |\boldsymbol{M}_n||\boldsymbol{B}_n^{\text{eff}}|$. Eq. (40), together with the angle ϕ chosen randomly in $[0, 2\pi)$ gives the value of the new direction of the moment $\boldsymbol{M}_n$ in the local system.

Since the heat-bath method has no rejection, it should at first sight be faster than the Metropolis method. However, it has two obvious drawbacks. The first is that the exponentials and the logarithm in Eq. (40) are numerically by far more expensive than the evaluation of $\cos\theta_n = 2\left(\frac{1}{2} - r\right)$ as a uniform random variable in the range $[-1, 1)$ that is required in the Metropolis method. The second drawback is that the method cannot be easily generalized, since it relies on the transformation of the random variable from the uniform to the new distribution. For the Heisenberg model this is given by Eq. (40), but for a different model an analytical expression might not be obtainable in closed form. For example, as soon as one adds a single-ion anisotropy term $-KM_z^2$ to the Heisenberg Hamiltonian, Eq. (40) is no more valid.

A further point to have in mind is the need for fast convergence of the stochastic process in terms of the number of Monte Carlo time-steps. This largely depends on the method, and it can be that for certain problems the Metropolis method is very efficient. Still, there is at least one problem where the heat-bath method for the Heisenberg model has proven to be by far superior to the Metropolis method: this is the problem of finding the Curie temperature in diluted magnetic semiconductors [21]. These systems comprise magnetic defects that are diluted in a non-magnetic, semiconducting or insulating host at a concentration of a few percent. Due to the band gap of the host, the exchange interaction between the impurities falls off exponentially with distance, resulting in a situation that most magnetic ions are only weakly coupled to their far-away neighbours, while some others that happen to be clustered together are strongly bound to their nearest magnetic neighbours. As it turns out, under these circumstances case the Metropolis method has a very long autocorrelation time, exceeding by two or three orders of magnitude the autocorrelation time of the heat-bath method [22] that proves to be an extremely valuable tool.

Cluster-flip methods

We close this section with a brief mention of cluster-flip methods. In such methods, designed for the Ising or Heisenberg model, the stochastic rule does not process one magnetic moment at a

time but a whole cluster of moments. The algorithms start at a certain site and then add on some of its neighbours, then some of the added-neighbours' neighbours, etc., each with a certain probability that is depends on the temperature, on the relative moment direction, and on the interaction strength. Once the addition of neighbours is exhausted, a connected cluster has been formed whose magnetic moments are flipped as a whole. One of the benefits of such methods is that they approach the configurations of the critical point by a physically motivated algorithm, since close to the critical point the system shows long-range fluctuations. Two well-known cluster-flip methods are the Wolff algorithm [23] and the Swendsen-Wang algorithm [24], both of which are well discussed by Newman and Barkema [18].

5 Application to bcc Fe

We come now to an application of the theory in the case of bcc Fe, one of the few elemental ferromagnets. We first summarize a few technical details of the calculation. The electronic structure calculations were carried out within the generalized gradient approximation to density-functional theory and employed the full-potential Korringa-Kohn-Rostoker Green function method [25]. Relativistic effects were taken into account within the scalar relativistic approximation. An angular momentum truncation at $l_{\max} = 3$ was taken. The exchange constants for the Heisenberg Hamiltonian (1) were calculated using the method of infinitesimal rotations that was discussed earlier here (Eq. (16)) and that was derived for the Korringa-Kohn-Rostoker method by Liechtenstein et al. [8]. For the calculation of thermodynamic properties, the Monte Carlo method was used, employing the Metropolis algorithm in periodic supercells. The number of sampling steps varied depending on the temperature, but close to $T_{\rm c}$ each atom in the supercell was sampled 8000 times at each temperature. As a random number generator we used the Mersene Twister code [26].

We continue by discussing the calculated exchange interactions. Fig. 1 shows the spin-resolved density of states at the ferromagnetic state together with the nearest-neighbour $J_{nn'}(E)$, that we call $J_1(E)$, given by Eq. (18). At $E_{\rm F}$, $J_1(E_{\rm F})$ is the physical interaction. We see in Fig. 1 that $J_1(E)$ obtains positive and negative values, i.e., there are ferromagnetic and antiferromagnetic contributions from the various bands. This has an origin in the bonding character between the states of nearest neighbours. We provide two examples, while we refer to Ref. [21] for a more thorough discussion. When rotating the two moments with respect to each other, some peak of the density of states centered at $E_{\rm F}$ may reduce its width because the same-spin hybridization of neighbouring wavefunctions lessens. This gives a ferromagnetic contribution to $J_1(E_{\rm F})$, because the reduced broadening means that occupied states are driven to higher energies, increasing the total energy by rotation. In a second example, two peaks of opposite spin might be centered at $E_{\rm F} + \Delta E$ for spin up and $E_{\rm F} - \Delta E$ for spin down. In the ferromagnetic state these states are orthogonal by spin and do not hybridize. But when the moments rotate, then there is a non-vanishing projection of the spin-up states of one atom to the spin-down of the other, and this results in hybridization and mutual repulsion of the peaks. Since the occupied peak will be shifted lower by this repulsion, a reduction of energy occurs by rotation, indicating an antiferromagnetic contribution. The same principle holds for the whole curve $J_1(E)$. The sum of all these contributions yields the displayed curve.

Next we examine the parameters $J_{nn'}$ as a function of distance between n and n'. These are shown in Fig. 2 (red dashed line). Since it obviously matters not only how strong is the interaction at a certain distance, but also how many equivalent neighbours there are in a shell of atoms

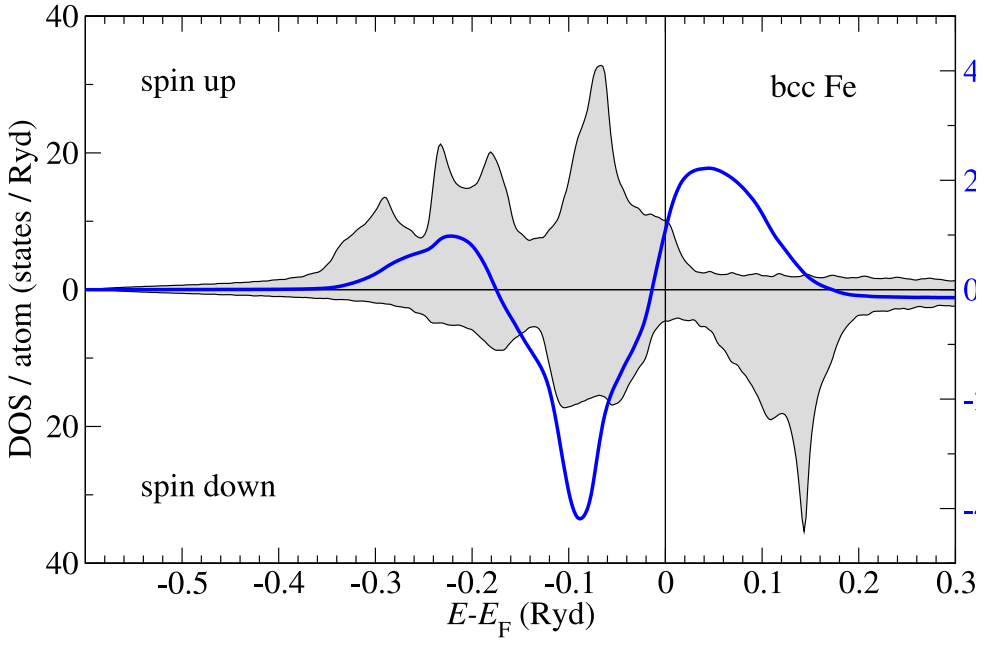


Fig. 1: *Spin-resolved density of states (DOS) (shaded area, left-hand scale) and nearest neighbour exchange interaction as a function of energy according to Eq. (18) (blue, right-hand scale) for bcc Fe.*

at this distance, the parameters $J_{nn'}$ have been multiplied by the number of neighbours in the shell and are depicted by the full green line. The strongest interactions are the ones to the nearest and second-nearest neighbours, where there is a direct overlap of d wavefunctions. At higher distances we clearly witness an oscillatory behaviour that is a Fermi-surface effect. These oscillations indicate that many shells of neighbours are necessary to obtain a convergence in the thermodynamic values. Indeed, when calculating the T_c either by mean-field theory (Eq. (19)) or by the Monte Carlo method (discussed below), we are practically forced to set a cutoff in the interaction distance, dropping the $J_{nn'}$ father than this cutoff. Fig. 2 shows the calculated T_c as a function of the cutoff within mean-field theory (upper blue curve) and Monte Carlo (lower blue curve).[10] The value of T_c is evidently oscillating with distance just as the $J_{nn'}$ with rather slow convergence. The experimental value of $T_\mathrm{c} = 1043$ Kelvin is underestimated by approximately 17% by the calculation that gives 870 Kelvin. This underestimation is not universal, since in other materials it can be larger or smaller or also an an overestimation can occur.

Now we come to the Monte Carlo calculations. The calculated quantities from the Monte Carlo simulation are the magnetization $\langle m \rangle(T)$, the (longitudinal) susceptibility per atom χ_T (Eq. (34)), the specific heat per atom $c_T = C_T/N_\mathrm{at}$ according to Eq. (26), and the fourth-order cumulant,

$$U_4(T) = 1 - \frac{1}{3}\frac{\langle m^4 \rangle}{\langle m^2 \rangle^2} \tag{41}$$

The role of the latter quantity will be discussed later on.

The Monte Carlo results are presented in Fig. 3. We show calculations for different supercell sizes to demonstrate the finite-size effects. The distance cutoff was taken at 4.5 lattice parameters.

[10] We remind the reader that mean-field theory always overestimates the Curie temperature.

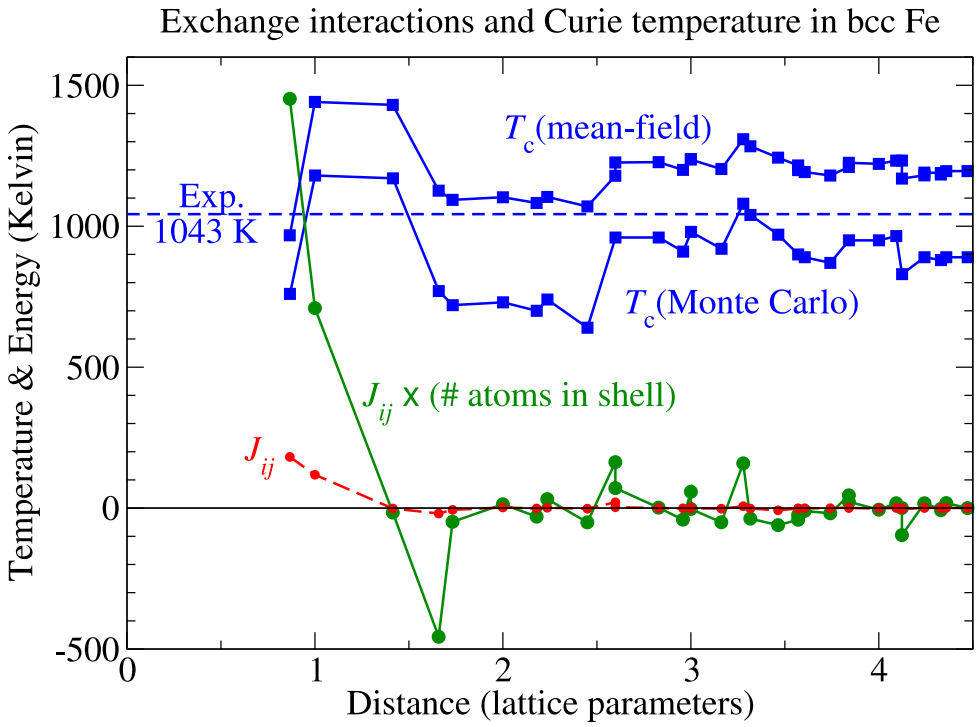


Fig. 2: *Variation of calculated Curie temperature of bcc Fe as a function of the cutoff in the distance-dependent exchange parameters. Shown is the mean-field and Monte Carlo derived* T_c *in a* $12\times12\times12$ *supercell, the exchange parameters* $J\,M^2$ *(red dashed curve) and the exchange parameters multiplied by the number of atoms in each. The units of* $J\,M^2$ *are transformed to Kelvin by multiplying with* k_B.

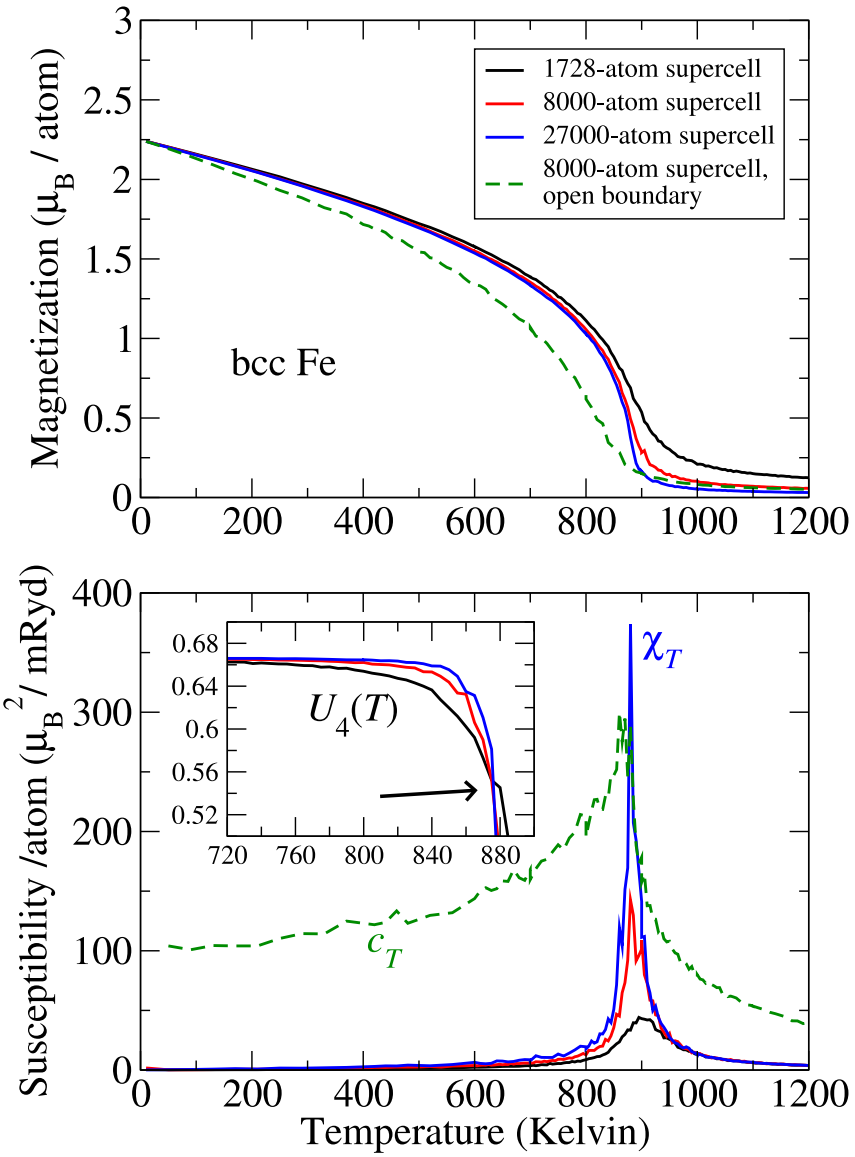


Fig. 3: *Calculated magnetization curve (upper panel) and susceptibility per atom χ_T (lower panel) for bcc Fe using the Monte Carlo method for three different periodic supercell sizes, $12\times12\times12$, $20\times20\times20$, and $30\times30\times30$. For larger cells the magnetization curve falls more abruptly and the susceptibility becomes more peaked as T approaches T_c. The magnetization curve for a $20\times20\times20$ supercell with open boundary conditions is also depicted, demonstrating the underestimation of T_c in this case due to surface effects. The inset in the lower panel shows the cumulant $U_4(T)$ for the case of the three supercell sizes. The crossing point of the cumulants, shown by an arrow, converges fast with supercell size towards the value of T_c. Also the specific heat per atom, c_T/k_B, for the $30\times30\times30$ supercell is depicted in the lower panel with a dashed curve (multiplied by a factor 100 to agree with the scale). The fluctuations in the curves are highest close to T_c and can be reduced by increasing number of sampling steps. Here, close to T_c, the supercell was sampled 8000 times for each temperature using the Metropolis method.*

At $T = 0$ the system is at the ferromagnetic state, the magnetization has its highest value, and the longitudinal susceptibility vanishes. This is understandable, since the moments cannot be elongated, by assumption of the model. The specific heat, however, does not vanish, since the system has the possibility to be excited to infinitesimally small energies absorbing energy even at $T = 0$. Here we see a clear difference to the quantum model, where the presence of a gap between the ground-state energy and the first excited level causes the specific heat to vanish. As the temperature increases, the individual moments deviate in direction from the average, thus they can be re-oriented by a field that is along the average moment. This causes χ_T to slowly rise. The magnetization drops linearly at first, as was discussed in the last paragraph of Section 2.

In the vicinity of the Curie temperature the magnetization drops fast and almost reaches zero. This is actually one way of locating T_c. The fact that $\langle m \rangle(T)$ does not exactly vanish above T_c is a finite-size effect of the simulation supercell, as is demonstrated by the different curves for different supercell sizes. Clearly the curve for the 27000-atom supercell (Fig. 3, upper panel) is much closer to what one expects from a second-order transition, i.e., that $\langle m \rangle(T)$ should drop to zero continuously but with discontinuous derivative at T_c. Next to these curves we also see the magnetization of an 8000-atom supercell but with open boundary conditions (dashed line); compared to the same-size supercell with periodic boundary conditions, the former clearly gives a lower value for T_c. Intuitively, one expects that any surface effects of the open boundary conditions should disappear, as the supercell becomes larger. However, this happens too slowly in the Heisenberg model. Thus, periodic boundary conditions improve the quality of the results.

The susceptibility, depicted in the lower panel of Fig. 3, shows a peak at T_c. The peak becomes sharper and taller as the supercell size increases, and it is known that in the limit of $N_\mathrm{at} \to \infty$ the susceptibility shows a divergence at T_c. Actually the position of the peak gives a very good estimate for T_c already at moderate supercell sizes. In order to understand this behaviour better, consider that the state at T_c comprises fluctuations at all length scales (as is typical for a second order phase transition). It includes large volumes, of arbitrarily large size (limited only by the supercell size in the calculation and by the sample size in reality) that may appear almost collinear but with the magnetization inside them pointing in any arbitrary direction. These volumes, say of size Ω interact with their surrounding moments via their surface only. Then, by applying a field, and as long as they keep their mutual angles, their reorientation as a whole costs energy proportional to their surface only, i.e. grows as $\Omega^{2/3}$, while the energy gain is proportional to Ω. At large Ω the surface energy cost becomes insignificant next to the volume energy gain, so that the reorientation of the magnetization in Ω towards the field direction is complete and the susceptibility diverges. From another perspective, these volumes are not static; they change in size, position, and orientation of magnetization as a function of time, causing extremely large fluctuations in time along with the large fluctuations in space. Since the susceptibility is proportional to the squared fluctuation amplitude (Eq. (34)) it diverges along with the fluctuations.

The same is true for the specific heat, as it depends on the energy fluctuation (Eq. (26)). At T_c the fluctuations of the energy are also diverging, and along with them follows c_T. Thus the peak of c_T is also a good estimate for the determination of T_c or of a phase transition in any system, even if the order parameter is not known.

From the lower panel of Fig. 3 we see that the susceptibility peak occurs at ever so slightly lower temperature as the supercell size increases. The limit of T_c for $N_\mathrm{at} \to \infty$ is found by different procedures that are collectively named *finite size scaling*. In the ferromagnetic Heisenberg model, one very powerful way for finite size scaling is to examine the fourth-order

cumulant $U_4(T)$ that was defined in Eq. (41) (see, e.g., Landau and Binder [19]). This quantity is monotonically decreasing, takes the values $U_4(0) = \frac{2}{3}$ and $U(\infty) = 0$, and for the infinite-size supercell limit it is just a step-function at T_c. At finite sizes, the values of $U_4(T)$ for different sizes cross at a point very close to T_c. The crossing point converges to T_c rapidly as the size increases. Thus one can perform the calculation of U_4 at two cells of moderate but different size, find the crossing point, and perhaps cross-check with a third cell.
For our example of bcc Fe the values of U_4 near T_c are shown in the inset of Fig. 3. The crossing of all three at almost the same point is evident, but if the statistical sampling was better and the curves were smoother, then we would see that the three curves do not cross at a single point; in that case the crossing point for the larger cells would be a better approximation to T_c. Here we should note that the necessary number of sampling steps to obtain well-converged averages increases with the power of the order parameter that needs to be sampled: we need more steps for $\langle m^2 \rangle$ than for $\langle m \rangle$ and even more for $\langle m^4 \rangle$. This is why the susceptibility values are fluctuating stronger than the magnetization values in Fig. 3.

6 Epilogue: Critical slowing down

We close this chapter with a comment on numerical accuracy, in particular on the necessary number of sampling steps close to T_c. This number increases as one approaches T_c, where the fluctuations become large. It increases still more as the supercell becomes larger, merely because then the fluctuation amplitude becomes larger (proportional to the number of atoms in the supercell at T_c). This effect of second-order phase transitions is known as *critical slowing down* and hampers most Monte Carlo methods [18, 19]. It has, however, also a physical origin, since in an experiment that is performed at T_c the fluctuations would also be large and slowly changing, and one would need a longer observation time to measure averaged quantities, compared the time at temperatures lower or higher than T_c. Normally the time scale involved is anyhow very short compared to the observation time, but new ultrafast probe techniques in magnetism can reach down to the femtosecond time-scale, i.e., far below the time-scale of the fluctuations. The effect of critical slowing down is also observed in real-life situations, especially in the scientific task of completing and polishing a manuscript or a thesis. Even when one naively believes that little is left to be done, critical slowing down makes the last corrections and adjustments take forever. In the same spirit as in phase-transition simulations, one is forced in real life to accept a cutoff for the time of optimization, that is otherwise only limited by the natural cutoff of human lifetime.

References

[1] M. V. You and V. Heine, *Magnetism in transition metals at finite temperatures. I. Computational model*, J. Phys. F: Met. Phys. **12**, 177 (1982).

[2] J. Staunton, B. L. Gyorffy, A. J. Pindor, G. M. Stocks, and H. Winter, *Electronic structure of metallic ferromagnets above the Curie temperature*, J. Phys. F: Met. Phys. **15**, 1387 (1985).

[3] V. P. Antropov, M. I. Katsnelson, B. N. Harmon, M. van Schilfgaarde, and D. Kusnezov, *Spin dynamics in magnets: Equation of motion and nite temperature effecs*, Phys. Rev. B **54**, 1019 (1996).

[4] S. V. Halilov, H. Eschrig, A. Y. Perlov, and P. M. Oppeneer, *Adiabatic spin dynamics from spin-density-functional theory: Application to Fe, Co, and Ni*, Phys. Rev. B **58**, 293 (1998).

[5] R. Gebauer and S. Baroni, *Magnons in real materials from density-functional theory*, Phys. Rev. B **61**, R6459 (2000).

[6] E. Sasioglu, L. M. Sandratskii, P. Bruno, and I. Galanakis, *Exchange interactions and temperature dependence of magnetization in half-metallic Heusler alloys*, Phys. Rev. B **72**, 184415 (2005).

[7] P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai, *Ground States of Constrained Systems: Application to Cerium Impurities*, Phys. Rev. Lett. **53**, 2512 (1984).

[8] A. I. Liechtenstein, M. I. Katsnelson, V. P. Antropov, and V. A. Gubanov, *Local spin density functional approach to the theory of exchange interactions in ferromagnetic metals and alloys*, J. Magn. Magn. Mater. **67**, 65 (1987).

[9] I. Turek, J. Kudrnovsky, V. Drchal, and P. Bruno, *Exchange interactions, spin waves, and transition temperatures in itinerant magnets*, Philosoph. Mag. **86**, 1713 (2006).

[10] A. Jacobsson, B. Sanyal, M. Ležaić, and S. Blügel *Exchange parameters and adiabatic magnon energies from spin-spiral calculations*, Phys. Rev. B **88**, 134427 (2013).

[11] A nice description of mean-field theory for the case of one or more than one magnetic atom in the unit cell is given by Philip W. Anderson, *Theory of magnetic exchange interactions: exchange in insulators and semiconductors*, in Solid State Physics, edited by F. Seitz and T. Turnbull, Vol. 14, p. 99 (Academic Press 1963).

[12] S. V. Tyablikov, *Methods of Quantum Theory of Magnetism* (Plenum Press, 1967).

[13] H. B. Callen, *Green Function Theory of Ferromagnetism*, Phys. Rev. **130**, 890 (1963).

[14] P. Bruno, *Exchange Interaction Parameters and Adiabatic Spin-Wave Spectra of Ferromagnets: A Renormalized Magnetic Force Theorem*, Phys. Rev. Lett. **90**, 087205 (2003).

[15] A. R. Mackintosh and O. K. Andersen, in *Electrons at the Fermi Surface*, edited by M. Springford , p. 149 (Cambridge University Press, 1980).

[16] A. Oswald, R. Zeller, P. J. Braspenning, and P. H. Dederichs, *Interaction of magnetic impurities in Cu and Ag*, J. Phys. F: Met. Phys. **15**, 193 (1985).

[17] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *Equation of State Calculations by Fast Computing Machines*, J. Chem. Phys. **21**, 1087 (1953).

[18] M. E. J. Newman and G. T. Barkema, *Monte Carlo methods in statistical physics* (Oxford University Press, 1999).

[19] David P. Landau and Kurt Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, 2000).

[20] Y. Miyatake, M. Yamamoto, J. J. Kim, M. Toyonaga, and O Nagai, *On the implementation of the 'heat bath' algorithms for Monte Carlo simulations of classical Heisenberg spin systems'* J. Phys. C: Solid State Phys. **19**, 2539 (1986).

[21] K. Sato, L. Bergqvist, J. Kudrnovský, P. H. Dederichs, O. Eriksson, I. Turek, B. Sanyal, G. Bouzerar, H. Katayama-Yoshida, V. A. Dinh, T. Fukushima, H. Kizaki, and R. Zeller, *First-principles theory of dilute magnetic semiconductors*, Rev. Mod. Phys. **82**, 1633 (2010).

[22] Brahim Belhadji, *Ab-initio Calculations for Dilute Magnetic Semiconductors*, PhD Thesis, RWTH Aachen University (2008).

[23] U. Wolff, *Collective Monte Carlo Updating for Spin Systems*, Phys. Rev. Lett. **62**, 361 (1989).

[24] R. H. Swendsen and J.-S. Wang, *Nonuniversal critical dynamics in Monte Carlo simulations*, Phys. Rev. Lett. 58, **86** (1987).

[25] The Korringa-Kohn-Rostoker method is based on a multiple-scattering expansion of the Green function. See the website `http://www.kkr-gf.org` for monographs or review articles on the method. The electronic structure calculations here employed *the SPR-TB-KKR Package*, H. Ebert and R. Zeller, `http://olymp.cup.uni-muenchen.de/ak/ebert/SPR-TB-KKR`.

[26] M. Matsumoto and T. Nishimura, *Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator*, ACM Trans. Model. Comput. Simul. **8**, 3 (1998).

C 2 Density Functional Perturbation Theory[1]

M.J. Verstraete and Z. Zanolli
Department of Physics and Peter Grünberg Institut
University of Liège, Belgium and
Forschungszentrum Jülich GmbH, Germany

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Density Functional Theory has emerged as a very successful technique to calculate the ground state properties of electronic systems, whether in bulk, nanostructures, or molecules. Formally DFT only provides the ground state density and total energy, and there are well known failures in the representation of electronic excited states. If however the ground state is perturbed very lightly, typically in such a manner as to leave the system close to its electronic ground state, one can easily apply the traditional quantum formalism for perturbation theory, giving density functional perturbation theory (DFPT). This approach was invented at the end of the 1980s, and has been particularly successful in predicting many experimentally observable quantities: experiment always probes excitations of a system, but for low energy "slow" excitations DFPT builds on the very good representation of the ground state in DFT to give accurate excitations as well. Higher energies imply more electronic transitions, in which the intrinsic DFT is a much worse approximation. The cures to these woes (TDDFT, GW etc...) are presented in several other chapters of these proceedings.

The first development of DFPT came from Baroni, Giannozzi, and Testa in Trieste [1]. Several groups followed suite in the 1990s, in particular those of Xavier Gonze, Serguei Savrasov, Amy Liu, and then many others. A large number of electronic structure simulation packages now implement and use DFPT. The historic ones are Quantum Espresso [2] and ABINIT [3], but also VASP [4], octopus [5], CASTEP [6] and many others. DFPT has become a popular tool for the analysis of the vibrational and spectroscopic properties of materials, routinely providing phonon frequencies and dielectric properties with just a few percent of error with respect to experiment.

In the following sections we will give basic definitions and review the important concepts of DFPT. Some derivations will be followed through, but the details will be left buried in the appropriate literature. The aim is to provide a framework and support for the lectures in the FZJ Spring School 2014, and to showcase a variety of representative and interesting applications which have links to other Lectures from the school. By the end of this Lecture students should have an accurate idea of what they can calculate with DFPT, how the calculation is carried out, and which experimentally observable quantities can be extracted. A number of review articles have been published over the past 25 years (e.g. Ref. [7] which we referred to frequently), in particular the Review of Modern Physics by Baroni et al. [8].

This chapter is dedicated to the memory of Professor Jean-Pol Vigneron (1950-2013).

2 Density Functional Perturbation Theory

2.1 Response functions

Many interesting physical properties (*response functions*) are the result of the application of an external perturbation to the system under investigation. Response functions are second, third, or higher order derivatives of the total energy with respect to applied perturbation(s). Typical perturbations can be atomic displacements, homogeneous external electric or magnetic fields, strain, alchemical change, etc... The physical properties related to the derivatives of the total (electronic plus ionic) energy are

1^{st} **order:** forces, stress, dipole moment, ...

2^{nd} **order:** phonon dynamical matrix, elastic constants, dielectric susceptibility, Born effective charges, piezoelectricity, internal strain

3^{rd} **order:** non-linear dielectric susceptibility, phonon-phonon interaction, Grüneisen parameters, anharmonic elastic constants, ...

Further physical properties - such as entropy or thermal expansion - can be obtained integrating the total energy (or a thermodynamic potential) over phononic degrees of freedom.
Total energy derivatives can be computed using either direct approaches (e.g., finite differences, molecular-dynamics spectral analysis) or perturbative techniques. The former suffer from various limitations such as size effects, ergodicity constraints, commensurability problems or the difficulty to decouple the responses to perturbations of different wavelength. The perturbative theory applied to DFT, instead, allows one to treat not only periodic perturbations but also perturbations characterized by a non-zero, commensurate or incommensurate, wavevector [9]. In addition, the computation of 1^{st} order corrections to wavefunctions within perturbation theory can be done using a variational approach (Section 2.3) and algorithms similar to those used for ground-state (i.e. unperturbed) calculations. Hence, the perturbative technique can be naturally included in the usual DFT framework. In the following, unless otherwise specified, we will use Hartree atomic units $\hbar = m_e = e = 1/4\pi\varepsilon_0 = 1$

2.2 Perturbation theory

Any perturbation theory starts by the identification of a small parameter, $\lambda << 1$, characterizing the change in the unperturbed Hamiltonian of the system $H^{(0)}$ due to an external potential $V_{ext}(\lambda)$:

$$H(\lambda) = H^{(0)} + V_{ext}(\lambda) \tag{1}$$

We assume that the unperturbed Schrödinger equation

$$H^{(0)} \left|\psi_i^{(0)}\right\rangle = \epsilon_i^{(0)} \left|\psi_i^{(0)}\right\rangle \tag{2}$$

can be solved exactly and the normalization condition on the unperturbed wavefunctions is satisfied

$$\left\langle \psi_i^{(0)} \middle| \psi_i^{(0)} \right\rangle = 1 \tag{3}$$

where i labels eigenvalues and eigenvectors of $H^{(0)}$. Our goal is to solve the perturbed Schrödinger equation

$$H(\lambda) \left|\psi_i(\lambda)\right\rangle = \epsilon_i(\lambda) \left|\psi_i(\lambda)\right\rangle \tag{4}$$

with the normalization condition on wavefunctions

$$\langle \psi_i(\lambda) \,|\, \psi_i(\lambda) \rangle = 1 \tag{5}$$

The next section will describe the specific types of perturbations usually examined. The fundamental hypothesis is that all of the observables and physical quantities can be expanded in Taylor series with respect to λ, and that these series are well defined and will converge. This constraint is not trivial formally, but in practice is usually verified. A perturbation of the Hamiltonian will change the resulting eigenstates and eigenvalues, generically to arbitrary order in λ. Let

$$H(\lambda) = H^{(0)} + \lambda H^{(1)} + \lambda^2 H^{(2)} + ... \tag{6}$$

Then the resulting eigensolutions for the full H are:

$$\psi_i(\lambda) = \psi_i^{(0)} + \lambda \psi_i^{(1)} + \lambda^2 \psi_i^{(2)} + ... \tag{7}$$

and eigenvalues

$$\epsilon_i(\lambda) = \epsilon_i^{(0)} + \lambda \epsilon_i^{(1)} + \lambda^2 \epsilon_i^{(2)} + ... \tag{8}$$

The n^{th} order correction to the eigenvalues (respectively, eigenvectors) is n^{th}-derivative of the eigenvalue (respectively, eigenvectors) computed at $\lambda = 0$. For instance

$$\epsilon_i^{(1)} = \left. \frac{d\epsilon_i(\lambda)}{d\lambda} \right|_{\lambda=0} \tag{9}$$

Hence the utility of perturbation theory to compute 1^{st}, 2^{nd} and 3^{rd} order derivatives of the total energy with respect to small perturbations. For certain perturbations the zeroth or first order, or, e.g., all even order corrections can be 0, depending on symmetry. In Eq. 7 we will presume that the eigenstate starts from a Kohn Sham (KS) independent particle orbital, and that $H^{(0)}$ is the corresponding KS Hamiltonian (the same form would be true for the full Many Body wavefunction and hamiltonian, of course). The new state is again a stationary solution of the new Hamiltonian, and we restrict ourselves to static perturbations for the moment. Replacing the above in the eigenequation for ψ (Eq. 4) and identifying terms in powers of λ gives a hierarchical and constructive set of equations for progressively higher order solutions:

$$\begin{aligned} H^{(0)} \,|\psi_i^{(1)}\rangle + H^{(1)} \,|\psi_i^{(0)}\rangle &= \epsilon_i^{(0)} \,|\psi_i^{(1)}\rangle + \epsilon_i^{(1)} \,|\psi_i^{(0)}\rangle & (10) \\ H^{(0)} \,|\psi_i^{(2)}\rangle + H^{(1)} \,|\psi_i^{(1)}\rangle + H^{(2)} \,|\psi_i^{(0)}\rangle &= \epsilon_i^{(0)} \,|\psi_i^{(2)}\rangle + \epsilon_i^{(1)} \,|\psi_i^{(1)}\rangle + \epsilon_i^{(2)} \,|\psi_i^{(0)}\rangle & (11) \end{aligned}$$

...

with the 0^{th} order Schrödinger equation given by Eq. 2. Inserting the perturbative series of ψ (7) in the normalization condition (5) we obtain a condition on the orthonormality of wavefunctions at each order:

$$\begin{aligned} \left\langle \psi_i^{(0)} \middle| \psi_i^{(1)} \right\rangle + \left\langle \psi_i^{(1)} \middle| \psi_i^{(0)} \right\rangle &= 0 & (12) \\ \left\langle \psi_i^{(0)} \middle| \psi_i^{(2)} \right\rangle + \left\langle \psi_i^{(1)} \middle| \psi_i^{(1)} \right\rangle + \left\langle \psi_i^{(2)} \middle| \psi_i^{(0)} \right\rangle &= 0 & (13) \end{aligned}$$

...

where the 0^{th} order normalization condition is given by Eq. 3. Left multiplying the 1^{st} order Schrödinger equation (10) by $\langle \psi_i^{(0)}|$ and using the orthonormality conditions (3) and (12) we obtain the 1^{st} order eigenvalues:

$$\epsilon_i^{(1)} = \langle \psi_i^{(0)}| \, H^{(1)} \,|\psi_i^{(0)}\rangle \tag{14}$$

This result is equivalent to the Hellmann-Feynman theorem [10, 11], which states that the first derivative of the eigenvalues of the Hamiltonian $H(\lambda)$ is given by the expectation value of the derivative of the Hamiltonian with respect to the parameter λ. From Eq. 14 it follows that the first-order corrections to the energy can be computed from the unperturbed wavefunctions and from the first-order change in the external potential. Higher-order terms are not needed to compute $\epsilon_i^{(1)}$.
The first-order correction to the wavefunctions can be computed by solving the so-called Sternheimer equation [12]:

$$(H^{(0)} - \epsilon_i^{(0)})\,|\psi_i^{(1)}\rangle = -(H^{(1)} - \epsilon_i^{(1)})\,|\psi_i^{(0)}\rangle \tag{15}$$

which is obtained by rearranging the terms in the 1^{st} order Schrödinger equation (10). Solving the Sternheimer equation for the $\psi_i^{(0)}$ requires the inversion of the $(H^{(0)} - \epsilon_i^{(0)})$ operator. However, the inversion cannot be done for arbitrary vectors since $\epsilon_i^{(0)}$ is an eigenvalue of $H^{(0)}$. The problem can be solved by expanding $|\psi_i^{(1)}\rangle$ in the basis of $|\psi_i^{(0)}\rangle$ (which is orthonormal and complete):

$$|\psi_i^{(1)}\rangle = \sum_j c_{ij}^{(1)}\,|\psi_i^{(0)}\rangle \tag{16}$$

This allows one to separate the 0^{th} order wavefunctions in two subsets: the subset I of the wavefunction(s) corresponding to the eigenvector $\epsilon_i^{(0)}$ and the subset of the wavefunctions in the orthogonal subspace $I^{\perp}$. After some algebra, it is possible to determine the coefficient $c_{ij}^{(1)}$ with $j \in I^{\perp}$, which permits to write the 1^{st} order correction to the wavefunctions in terms of a sum over the spectrum of the unperturbed Hamiltonian:

$$|\psi_i^{(1)}\rangle = \sum_{j \in I^{\perp}} |\psi_j^{(0)}\rangle\,\frac{\langle\psi_j^{(0)}|\,H^{(1)}\,|\psi_i^{(0)}\rangle}{\epsilon_i^{(0)} - \epsilon_j^{(0)}} \tag{17}$$

This sum runs over all the states (both occupied and empty) of the system except the specific eigenstate under consideration, for which the denominator would vanish. The coefficient $c_{ij}^{(1)}$ with $j \in I$ remains unknown in this approach, which gives us a gauge freedom. Imposing the normalization condition on the 1^{st} order eigenvector results in $c_{ij}^{(1)} = 0$ when $j \in I$.
The Sternheimer equation can also be written in terms of the projector onto the $I^{\perp}$ subspace $P_{I^{\perp}} = \sum_{j \in I^{\perp}} |\psi_j^{(0)}\rangle\,\langle\psi_j^{(0)}|$ as

$$P_{I^{\perp}}\left(H^{(0)} - \epsilon_i^{(0)}\right) P_{I^{\perp}}\,|\psi_i^{(1)}\rangle = -P_{I^{\perp}} H^{(1)}\,|\psi_i^{(0)}\rangle \tag{18}$$

where a full identity $\mathbf{I} = P_{I^{\perp}} + P_I$ has been inserted in the left hand side, and applied to the left. The advantage of this expression is that the singularity has been removed and, hence, the Sternheimer equation can be solved for $\psi_i^{(1)}$. Defining the Green's function in the subspace $I^{\perp}$ as

$$G_{I^{\perp}}(\epsilon) = \left[P_{I^{\perp}}\left(\epsilon - H^{(0)}\right) P_{I^{\perp}}\right]^{-1} \tag{19}$$

the equation to solve becomes

$$P_{I^{\perp}}\,|\psi_i^{(1)}\rangle = G_{I^{\perp}}(\epsilon_i^{(0)}) H^{(1)}\,|\psi_i^{(0)}\rangle \tag{20}$$

i.e. Green's function techniques can be used to solve the Sternheimer equation (15).
Several expressions exist for the second-order corrections to the energy. A simple expression can be obtained by left multiplying the 2^{nd} order Schrödinger equation (11) by $\langle\psi_i^{(0)}|$ and using the 0^{th} order equations (2) and (3):

$$\epsilon_i^{(2)} = \langle\psi_i^{(0)}|\, H^{(2)}\, |\psi_i^{(0)}\rangle + \langle\psi_i^{(0)}|\, H^{(1)} - \epsilon_i^{(1)}\, |\psi_i^{(1)}\rangle \tag{21}$$

summing this expression to its Hermitian conjugate and using the 1^{st} order normalization condition (12) we obtain:

$$\epsilon_i^{(2)} = \langle\psi_i^{(0)}|\, H^{(2)}\, |\psi_i^{(0)}\rangle + \frac{1}{2}\left(\langle\psi_i^{(0)}|\, H^{(1)}\, |\psi_i^{(1)}\rangle + \langle\psi_i^{(1)}|\, H^{(1)}\, |\psi_i^{(0)}\rangle\right) \tag{22}$$

The 2^{nd} order correction to the energy can also be obtained by minimizing with respect to $\psi_i^{(1)}$ the expression

$$\begin{aligned}\epsilon_i^{(2)} = \min_{\psi_i^{(1)}} \{\, &\langle\psi_i^{(0)}|\, H^{(2)}\, |\psi_i^{(0)}\rangle + \langle\psi_i^{(1)}|\, H^{(0)} - \epsilon_i^{(0)}\, |\psi_i^{(1)}\rangle + \\ &\langle\psi_i^{(0)}|\, H^{(1)} - \epsilon_i^{(1)}\, |\psi_i^{(1)}\rangle + \langle\psi_i^{(1)}|\, H^{(1)} - \epsilon_i^{(1)}\, |\psi_i^{(0)}\rangle\, \}\end{aligned} \tag{23}$$

under the normalization constraint Eq. 12. The expression enclosed by curly brackets is usually referred to as the *variational expression* of the 2^{nd} order correction to the energy. Similar variational expressions exist for all even orders. The number of terms to be calculated is larger, and the non-variational expressions (which exist for all orders) are often used for simplicity in other cases. They can often obtain the full set of $\epsilon^{(2)}$ without needing the full set of $\psi^{(1)}$.

2.3 Density Functional Perturbation Theory

The DFT ground-state energy of the electronic system is obtained by minimizing the functional of the electronic density:

$$E_{el}[\rho(\lambda)] = \sum_{i=1}^{N_e} \langle\psi_i(\lambda)|\, (T + V_{ext}(\lambda))\, |\psi_i(\lambda)\rangle + E_{Hxc}[\rho(\lambda)] \tag{24}$$

under the orthonormality constraint on the Kohn-Sham orbitals:

$$\langle\psi_i(\lambda)\,|\,\psi_j(\lambda)\rangle = \delta_{ij} \tag{25}$$

The sum in Eq. 24 is performed on all the occupied states, T is the kinetic energy operator, V_{ext} is the external potential (which includes the nuclear potential), E_{Hxc} is the Hartree and exchange-correlation energy functional. The electronic density is given by

$$\rho(\mathbf{r};\lambda) = \sum_{i=1}^{N_e} \psi_i^*(\mathbf{r};\lambda)\psi_i(\mathbf{r};\lambda) = \rho^{(0)}(\mathbf{r}) + \lambda\rho^{(1)}(\mathbf{r}) + \lambda^2\rho^{(2)}(\mathbf{r}) + ... \tag{26}$$

and the minimization can be performed using the Lagrange multiplier method. The canonical Euler-Lagrange equations can be cast in the form of a Schrödinger equation (4) with the Hamiltonian

$$H = T + V_{ext}(\lambda) + V_{Hxc}(\lambda) \tag{27}$$

where

$$V_{Hxc}(\mathbf{r};\lambda) = \frac{\partial E_{Hxc}[\rho(\lambda)]}{\partial \rho(\mathbf{r})} \tag{28}$$

Analogously to standard perturbation theory (Section 2.2), it is assumed that the unperturbed problem ($\lambda = 0$) can be solved exactly and the perturbing potential is known at all orders. As mentioned in the Introduction, the goal of Density Functional Perturbation Theory (DFPT) is to compute the derivatives of the DFT electronic energy with respect to different perturbations. The full derivation of the (many) DFPT equations can be found in [1, 13, 14, 15]. In order to describe the whole Born-Oppenheimer energy, the electrostatic repulsion between nuclei must be added to the electronic contribution. The nuclei-nuclei interaction energy can be obtained by treating the nuclei as classical point charges, so it does not present particular computational difficulties and hence we focus here on the electronic contribution only.
In standard perturbation theory it is shown that the first order correction to the electronic energy can be computed from the 0^{th} order wavefunctions and the perturbing potential at the 1^{st} order (14):

$$E_{el}^{(1)} = \sum_{i=1}^{N_e} \langle \psi_i^{(0)} | \, H^{(1)} \, | \psi_i^{(0)} \rangle = \sum_{i=1}^{N_e} \langle \psi_i^{(0)} | \, (T + V_{ext})^{(1)} \, | \psi_i^{(0)} \rangle + \left. \frac{dE_{Hxc}[\rho^{(0)}]}{d\lambda} \right|_{\lambda=0} \tag{29}$$

which is the Hellmann-Feynman theorem for DFT.
The constraint on the wavefunctions at 1^{st} order is given by Eq. 12. The 1^{st} order change in the wavefunctions can be obtained by solving directly the Sternheimer equation (15) using iterative techniques or a self-consistent approach analogous to the self-consistent solution of the unperturbed problem where the Kohn-Sham equation is replaced by Eq. 15. Sternheimer gives a system of N_e linearly coupled equations: this is formally much simpler than standard DFT, where the $E_{xc}[\rho]$ dependency is non-linear. The $\psi_i^{(1)}$ - solution of Eq. 15 - can also be computed using the *sum over states* expression (Eq. 17) which only requires the knowledge of eigenvalues and eigenvectors of the unperturbed Hamiltonian, and the perturbing potential at 1^{st} order.
As discussed in Section 2.2 the Sternheimer equation can also be solved using Green's function techniques (Eq. 18). For semiconductors $P_{I^\perp}$ becomes the projector upon the conduction states. For metals one should take explicitly into account that a finite density of states is present at the Fermi energy and hence the external perturbation can affect the occupation numbers. The linear response of metals has been treated by de Gironcoli [16]. The limits of validity of DFPT are dictated by the requirement that the external perturbation is small, i.e. $\lambda \, |\psi_i^{(1)}\rangle \ll |\psi_j^{(0)}\rangle$. Hence Eq. (17) implies $|\lambda \, \langle \psi_j^{(0)} | \, H^{(1)} \, | \psi_i^{(0)} \rangle \,| \ll |\epsilon_i^{(0)} - \epsilon_j^{(0)}|$, which means that the external perturbation must be small with respect to electronic excitations.
Once the 1^{st} order change of wavefunctions $\psi_i^{(1)}$ is known, the first-order change of density can be obtained by inserting the expansion (7) in Eq. 26:

$$\rho^{(1)}(\mathbf{r}) = \sum_{i=1}^{N_e} \left[\psi_i^{*(1)}(\mathbf{r}) \psi_i^{(0)}(\mathbf{r}) + \psi_i^{*(0)}(\mathbf{r}) \psi_i^{(1)}(\mathbf{r}) \right] \tag{30}$$

and, hence, the computation of the first-order change in the Hamiltonian can be done using [1]

$$H^{(1)} = T^{(1)} + V_{ext}^{(1)} + V_{Hxc}^{(1)} = T^{(1)} + V_{ext}^{(1)} + \int \left. \frac{\partial^2 E_{Hxc}}{\partial \rho(\mathbf{r}) \partial \rho(\mathbf{r}')} \right|_{\rho^{(0)}} \rho^{(1)}(\mathbf{r}') d\mathbf{r}' \tag{31}$$

The 2^{nd} order derivative of the electronic energy can be given by a non-variational expression that follows from Eq. 22:

$$E_{el}^{(2)} = \sum_{i=1}^{N_e} \langle \psi_i^{(0)} | \, H^{(2)} \, | \psi_i^{(0)} \rangle + \sum_{i=1}^{N_e} \langle \psi_i^{(0)} | \, H^{(1)} \, | \psi_i^{(1)} \rangle \tag{32}$$

A more accurate variational expression (analogous to Eq. 23 of standard perturbation theory) can be obtained by minimizing

$$E_{el}^{(2)} \left[\left\{ \psi_i^{(0)} \right\}, \left\{ \psi_i^{(1)} \right\} \right] = \sum_{i=1}^{N_e} \Big[\langle \psi_i^{(0)} | \, (T + V_{ext})^{(2)} \, | \psi_i^{(0)} \rangle + \langle \psi_i^{(1)} | \, (H - \epsilon_i)^{(0)} \, | \psi_i^{(1)} \rangle + \langle \psi_i^{(0)} | \, (T + V_{ext})^{(1)} \, | \psi_i^{(1)} \rangle + \langle \psi_i^{(1)} | \, (T + V_{ext})^{(1)} \, | \psi_i^{(0)} \rangle \Big] + \frac{1}{2} \int \int \frac{\partial^2 E_{Hxc}[\rho^{(0)}]}{\partial \rho(\mathbf{r}) \partial \rho(\mathbf{r}')} \rho^{(1)}(\mathbf{r}) \rho^{(1)}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' + \int \frac{d}{d\lambda} \frac{\partial E_{Hxc}[\rho^{(0)}]}{\partial \rho(\mathbf{r})} \Big|_{\lambda=0} \rho^{(1)}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \frac{d^2 E_{Hxc}[\rho^{(0)}]}{d\lambda^2} \Big|_{\lambda=0} \tag{33}$$

with respect to $\{\psi_i^{(1)}\}$ under the constraints $\left\langle \psi_i^{(0)} \,\middle|\, \psi_j^{(1)} \right\rangle = 0$ for all occupied states i and all j. Since $E_{el}^{(2)}$ is variational with respect to $\{\psi_i^{(1)}\}$, this leads to Euler-Lagrange equations which are equivalent to the self-consistent Sternheimer equation (Eq. 18).

Equations (32) and (33) explicitly show that the 2^{nd} order derivatives of the total energy are completely determined by the 1^{st} order derivatives of the wavefunctions. In both standard perturbation theory and DFPT it can be shown that the 3^{rd} order correction to the energies can be obtained by knowing the correction to the wavefunctions up to the 1^{st} order. These results are actually more general and are a consequence of the so-called "$2n+1$" theorem of quantum mechanics. This theorem states that the knowledge of the derivatives of the wavefunctions at order n^{th} permits the computation of the derivatives of the energy up to order $2n+1$. The theorem follows from the variational principle and has been demonstrated in DFT by Gonze and Vigneron [13]. The practical importance of the "$2n+1$" theorem is that it allows one to access 3^{rd} order derivatives of the total energy, e.g. 3^{rd} order effects (such as anharmonic phonon line widths, Raman scattering cross sections, nonlinear optical response) at the same computational cost as harmonic properties.

As a final cherry on the cake, once certain first-order wavefunctions are known, the non-variational expressions produced from Eq. 32 can be used to obtain many more mixed perturbations, at no extra cost. As an example see the Born effective charges below, or the off-diagonal terms of the dielectric susceptibility. The same holds true for third order ("2n+1") quantities: normally only two of the three $\psi^{(1)}$ are explicitly needed to get all of the mixed $E^{(3)}$ derivatives.

3 The most common perturbations

3.1 Phonons

The quantum mechanical modeling of bulk, molecules and nanostructured materials is usually performed in the *adiabatic approximation* of Born and Oppenheimer [17]. This approximation relies on the much greater mass of nuclei with respect to electrons. From this it follows that the nuclear motion is much slower than the electrons motion. Hence, it is possible to decouple nuclear (i.e. vibrational) and electronic degrees of freedom in the Hamiltonian and perform the computation of the total energy of the system in two steps. At first, the kinetic energy of the nuclei is supposed to be constant and subtracted from the total Hamiltonian. The resulting Born-Oppenheimer Hamiltonian H_{BO} depends parametrically upon the nuclear positions τ and describes the problem of the interacting electrons moving in the electrostatic field of nuclei at fixed positions:

$$H_{BO}(\tau) = -\frac{1}{2}\sum_i \frac{\partial^2}{\partial \mathbf{r}_i^2} + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i,\kappa}\frac{Z_\kappa}{|\mathbf{r}_i - \tau_\kappa|} + \frac{1}{2}\sum_{\kappa\neq\mu}\frac{Z_\kappa Z_\mu}{|\tau_\kappa - \tau_\mu|} \tag{34}$$

where $\mathbf{r}$ are the electronic coordinates, latin (i, j) and greek (κ, μ) letters respectively label electron and nuclei, and Z is the nuclear charge. In H_{BO} the kinetic energy of the electrons and the various electrostatic interaction terms (the repulsive electron-electron, the attactive electron-nucleus, and the repulsive nucleus-nucleus) can easily be identified. Eq. 34 can be solved, for instance, using DFT techniques and gives the ground-state Born-Oppenheimer energy of the system $E_{BO}(\tau)$. Then the kinetic energy of the nuclei is added, resulting in a Schrödinger equation for the nuclear motion which determines the dynamics of the lattice of the system

$$\left(-\frac{1}{2}\sum_\kappa \frac{1}{M_\kappa}\frac{\partial^2}{\partial \tau_\kappa^2} + E_{BO}(\tau)\right)\Psi(\tau) = E\Psi(\tau) \tag{35}$$

where E is the total energy and M_κ are the masses of the atoms.
Let us suppose that the system under investigation is periodic and, hence, can be described by the periodic repetition of a unit cell. Let the vector $\mathbf{R}$ identify the position of a periodic image of the unit cell with respect to the chosen origin, κ label a nucleus within the unit cell, and β label the cartesian directions. The nuclei are not fixed to their classical zero-temperature ideal positions but actually perform small displacements $u_{\mathbf{R}\kappa\beta}$ around their equilibrium positions $\tau_{\mathbf{R}\kappa\beta}$. Hence the Born-Oppenheimer energy of a crystal can be expanded in a Taylor series as a function of the nuclear displacements

$$E_{BO}(\mathbf{u}) = E^0_{BO} + \frac{1}{2}\sum_{\mathbf{R}\kappa\beta}\sum_{\mathbf{R}'\kappa'\beta'}\frac{\partial^2 E_{BO}}{\partial\tau_{\mathbf{R}\kappa\beta}\partial\tau_{\mathbf{R}'\kappa'\beta'}}u_{\mathbf{R}\kappa\beta}u_{\mathbf{R}'\kappa'\beta'} + ... \tag{36}$$

where E^0 is the minimum of the energy in the static approximation, when all the displacements are zero, and can be calculated within standard DFT. The 1^{st} order term in the series represents the forces acting on each nucleus, which vanish at the equilibrium when E_{BO} reaches a minimum and $F_{\mathbf{R}\kappa\beta} = \partial E_{BO}/\partial\tau_{\mathbf{R}\kappa\beta} = 0$. Hence the linear contribution in the expansion (36) disappears. The term containing the 2^{nd} derivative of the energy is called *harmonic* and the Taylor expansion truncated at the 2^{nd} order is referred to as the *harmonic approximation*. Anharmonic effects are described by higher order terms in Eq. 36 and will be discussed in Section 4.1.

When a nucleus is displaced from its equilibrium position a force appears to bring the atom back, according to the principle of virtual work. In the harmonic approximation (36), the force acting on a nucleus presents a linear dependence on the displacement of that nucleus:

$$F(\mathbf{R}'\kappa'\beta') = \frac{\partial E_{BO}}{\partial u_{\mathbf{R}\kappa\beta}} = -\sum_{\mathbf{R}'\kappa'\beta'} \frac{\partial^2 E_{BO}}{\partial \tau_{\mathbf{R}\kappa\beta}\partial\tau_{\mathbf{R}'\kappa'\beta'}} u_{\mathbf{R}'\kappa'\beta'} = -\sum_{\mathbf{R}'\kappa'\beta'} \Phi^{(2)}(\mathbf{R}\kappa\beta;\mathbf{R}'\kappa'\beta')u_{\mathbf{R}'\kappa'\beta'} \tag{37}$$

Equation (37) defines the matrix of *interatomic force constants* (IFCs), which is used to describe the force acting on a specific nucleus κ due to the displacement of another nucleus κ'.

In classical mechanics the lattice dynamics is described by the Newton equations of motion whose solution (in the harmonic approximation) is a superposition of the *normal modes of vibration* of the system. Taking into account the periodicity of the lattice, the normal modes can be written as Bloch states, i.e. as the product of a plane wave (which depends on the cell $\mathbf{R}$ in the direct lattice) and a lattice-periodic function

$$U_{\mathbf{R}m\mathbf{q}}(\kappa\beta) = e^{i\mathbf{q}\cdot\mathbf{R}}U_{m\mathbf{q}}(\kappa\beta) \tag{38}$$

where the wavevector $\mathbf{q}$ is a vector in the first Brillouin zone and characterizes the normal modes of vibration. The harmonic ansatz also implies a simple phase dependency on time $\exp(-i\omega t)$ to be added to Eq. 38 before insertion in Newton's law. In quantum mechanics the nuclear displacements $U_{\mathbf{R}m\mathbf{q}}(\kappa\beta)$ are quantized and known as *phonons*. Phonon eigendisplacements and their frequencies $\omega_{m\mathbf{q}}$ can be obtained by solving a generalized eigenvalue equation for the periodic part of the Bloch state

$$\sum_{\kappa'\beta'} \tilde{\Phi}^{(2)}(\kappa\beta;\kappa'\beta';\mathbf{q})U_{m\mathbf{q}}(\kappa'\beta') = M_\kappa\omega^2_{m\mathbf{q}}U_{m\mathbf{q}}(\kappa\beta) \tag{39}$$

where the *dynamical matrix* $\tilde{\Phi}^{(2)}(\kappa\beta;\kappa'\beta';\mathbf{q})$ is the Fourier transform of the IFC matrix

$$\tilde{\Phi}^{(2)}(\kappa\beta;\kappa'\beta';\mathbf{q}) = \sum_{\mathbf{R}'} \Phi^{(2)}(\mathbf{0}\kappa\beta;\mathbf{R}'\kappa'\beta')e^{i\mathbf{q}\cdot\mathbf{R}'} \tag{40}$$

and the eigendisplacements satisfy the normalization condition

$$\sum_{\kappa\beta} M_\kappa[U_{m\mathbf{q}}(\kappa\beta)]^* U_{m\mathbf{q}}(\kappa\beta) = 1 \tag{41}$$

If the system under investigation has N atoms in the unit cell, the dimension of the dynamical matrix is $3N\times 3N$ and the eigendisplacements are vectors of length $3N$ describing the displacements of the N nuclei along the 3 Cartesian directions. The resulting phonon band structure is not unbounded from above in energy, as is the case for electrons, but contains exactly $3N$ bands, with usual energies in the 1-10 meV range (up to several hundreds of meV for bonds involving Hydrogen).

The interatomic force constants and dynamical matrices are 2^{nd}-order derivatives of the Born Oppenheimer energy with respect to atomic displacements. Hence they can be accurately and efficiently computed using DFPT.

3.2 Periodic systems and incommensurate perturbations

The quantum mechanical modeling of systems under a periodic external potential can easily be accomplished thanks to the Bloch theorem. In this section we will discuss how to treat the case of a perturbation characterized by an incommensurate wavevector, which breaks the translational symmetry.
Bloch's theorem demonstrates that the wavefunction of a system subject to a potential with translational periodicity $\mathbf{R}$ in the direct lattice

$$V_{ext}^{(0)}(\mathbf{r}+\mathbf{R}) = V_{ext}^{(0)}(\mathbf{r}) \tag{42}$$

can be expressed as the product of a plane wave characterized by a vector $\mathbf{k}$ of the reciprocal lattice and a periodic function $u_{j\mathbf{k}}^{(0)}$ having the same periodicity as the external potential:

$$\psi_{j\mathbf{k}}^{(0)}(\mathbf{r}) = \frac{1}{N_{cell}\Omega_0} e^{i\mathbf{k}\cdot\mathbf{r}} u_{j\mathbf{k}}^{(0)}(\mathbf{r}) \tag{43}$$

where N_{cell} is the number of replicas of the unit cell included in the Born-von-Karman supercell, Ω_0 is the volume of the unit cell, and j is the electronic band index. The electron density has the same periodicity of the wavefunctions and is given by integrating the wavevectors $\mathbf{k}$ in the first Brillouin zone and summing j on the occupied bands

$$n^{(0)}(\mathbf{r}) = \frac{1}{(2\pi)^3} \int_{BZ} \sum_{j}^{occ} u_{j\mathbf{k}}^{(0)*}(\mathbf{r}) u_{j\mathbf{k}}^{(0)}(\mathbf{r}) d\mathbf{k} \tag{44}$$

Let us consider a perturbation to the periodic potential $V_{ext}^{(0)}$ which is characterized by an incommensurate wavevector $\mathbf{q}$ [13, 15, 18]: such that

$$V_{ext,\mathbf{q}}^{(1)}(\mathbf{r}+\mathbf{R}) = e^{i\mathbf{q}\cdot\mathbf{R}} V_{ext,\mathbf{q}}^{(1)}(\mathbf{r}) \tag{45}$$

As written, this perturbing potential is Hermitian only when $\mathbf{q}$ is half of a vector of the reciprocal lattice. For generic $\mathbf{q}$, the sum of this potential and its Hermitian conjugate, $V_{ext,\mathbf{q}}^{(1)} + V_{ext,-\mathbf{q}}^{(1)}$, should be used instead, such that the Hamiltonian is Hermitian. With this potential we have:

$$V_{ext,\mathbf{q}}^{(1)}(\mathbf{r}+\mathbf{R}) = e^{i\mathbf{q}\cdot\mathbf{R}} V_{ext,\mathbf{q}}^{(1)}(\mathbf{r}) + e^{-i\mathbf{q}\cdot\mathbf{R}} V_{ext,-\mathbf{q}}^{(1)}(\mathbf{r}) \tag{46}$$

If we apply a translation $\mathbf{R}$, the 1^{st} order wavefuction can be obtained from Eq. (17) of perturbation theory:

$$\psi_{j\mathbf{kq}}^{(1)}(\mathbf{r}+\mathbf{R}) = e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}} \psi_{j\mathbf{kq}}^{(1)}(\mathbf{r}) \tag{47}$$

where the 0^{th} order wavefunction Eq. (43) and the periodicity of $u_{j\mathbf{k}}^{(0)}(\mathbf{r})$ have been used. The 1^{st} order density at $(\mathbf{r}+\mathbf{R})$ is given by Eq. (30)

$$n_{\mathbf{q}}^{(1)}(\mathbf{r}+\mathbf{R}) = e^{i\mathbf{q}\cdot\mathbf{R}} n_{\mathbf{q}}^{(1)}(\mathbf{r}) \tag{48}$$

In order to restore the translational periodicity of the unperturbed system, we need to eliminate the phase factor $e^{i\mathbf{q}\cdot\mathbf{R}}$ in the 1^{st} order wavefunction and density (as with ordinary applications of Bloch's theorem). To this end we define the following functions:

$$u_{j\mathbf{kq}}^{(1)}(\mathbf{r}) = \sqrt{N_{cell}\Omega_0} e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{r}} \psi_{j\mathbf{kq}}^{(1)}(\mathbf{r}) \tag{49}$$

$$\overline{n}_{\mathbf{q}}^{(1)}(\mathbf{r}) = e^{-i\mathbf{q}\cdot\mathbf{r}} n_{\mathbf{q}}^{(1)}(\mathbf{r}) \tag{50}$$

whose periodicity in $\mathbf{R}$ can be easily tested using Eq. (47) and Eq. (48). The latter relationships allow one to write the 1^{st} order change in the electron density as

$$\overline{n}^{(1)}_{\mathbf{q}}(\mathbf{r}) = \frac{2}{(2\pi)^3}\int_{BZ}\sum_{j}^{occ} s u^{(0)*}_{j\mathbf{k}}(\mathbf{r}) u^{(1)}_{j\mathbf{k}}(\mathbf{r}) d\mathbf{k} \tag{51}$$

In this manner, the 1^{st} order problem can be addressed by solving equations which have the same translational symmetry as the unperturbed system.

Thermodynamical quantities

At usual temperatures for Humans, up to perhaps a thousand Kelvin, the main contributors to the variation of entropy and internal energy are the phonons. Electronic energies, of order eV, correspond to tens of thousands of degrees, and, whether in metals or high-T insulators, the number of thermally active electronic degrees of freedom (i.e. those close to the chemical potential) is quite low. Other degrees of freedom, e.g. magnetic ones or strong or weak nuclear forces, correspond to very small or very large energy scales, and are usually inactive at ambient conditions, either frozen out or saturated uniformly. The vibrational modes of a system thus give quantitative access to many thermodynamical observables. These are derived directly from the number of accessible states at a given energy, viz. the normalized density of phonon states

$$g(\omega) = \frac{(2\pi)^3}{3N\Omega_0}\sum_m \int_{\mathrm{BZ}} \delta(\omega - \omega_{\mathbf{q}m}) d\mathbf{q} \tag{52}$$

The internal energy and entropy are expressed simply using Bose-Einstein statistics with a dimensionless argument $x = \hbar\omega/k_B T$:

$$\begin{aligned} E^{ph} &= (3N)\int_0^{\omega_{\max}} \left(\frac{e^x+1}{e^x-1}\right)\frac{\hbar\omega}{2} g(\omega) d\omega \\ S^{ph} &= (3N)k_B \int_0^{\omega_{\max}} \left(\frac{xe^x}{e^x-1} - \ln(e^x-1)\right) g(\omega) d\omega \end{aligned}$$

which combined give the free energy

$$F^{ph} = (3N)k_B T \int_0^{\omega_{\max}} \ln\left(2\sinh\frac{x}{2}\right) g(\omega) d\omega$$

whose T derivative is the constant volume specific heat

$$C_V^{ph} = (3N)k_B \int_0^{\omega_{\max}} \left(\frac{x}{e^{x/2}-e^{-x/2}}\right)^2 g(\omega) d\omega$$

So far we have considered only constant volume quantities in the (N, V, T) ensemble. As we will see in Sec. 4.1 interatomic forces are not purely harmonic and this leads to variations of V with T - most experiments are performed at constant pressure. The simplest approach to this is the *quasi-harmonic* approximation: performing several calculations for different volumes, we can obtain both variations of F^{ph} with V and T and the full equation of state (EOS), presuming all anharmonic effects are in the volume expansion. Extracting the bulk modulus B and the volume thermal expansion coefficient α_V from the EOS, one can derive the constant pressure specific heat: $C_P^{ph} = C_V^{ph} + \alpha_V^2(T) B \Omega_0 T$

3.3 Macroscopic Electric Field

Overview and k derivatives

After atomic displacements, the next most common perturbation is that by a macroscopic electric field $\mathcal{E}$. Its inclusion in DFPT is relatively simple, in that no term in the normal BO Hamiltonian depends explicitly on the electric field, only implicitly through induced changes in the wave functions and density. Thermodynamically an external electric field couples directly to the electrical dipole moment or polarization $\mathcal{P}$, and we must add a term $\mathcal{E} \cdot \mathcal{P} \sim \int d\mathbf{r} \mathcal{E} \cdot \hat{r}$ to the Hamiltonian.
As is, this term is unbounded and ill defined in periodic boundary conditions. For the ground state it should be treated with the modern theory of polarization, see Ref. [19], Lecture A6, and comments below. This is possible and has other advantages, but makes the formalism more involved. Here for the description of linear responses it will be sufficient to start with a finite $\mathbf{q}$ field, whose potential is bounded, and whose wavelength will later be taken to infinity ($\mathbf{q} \to 0$). The chosen dependency on $\mathcal{E}$ limits us to longitudinal fields, but again this will be the response function we are interested in. Finally we will consider static fields and insulators - the response of metals is more complex (electrical transport is considered in Sect. 4.4).
In general one can relate the macroscopic electric field with an electrical polarization $\mathcal{P}$ and the corresponding displacement $\mathcal{D} = \mathcal{E} + 4\pi\mathcal{P}$. The linear term in $\mathcal{D}$ gives the (purely electronic part of the) dielectric permittivity tensor:

$$\varepsilon^{\infty}_{\alpha\beta} = \delta_{\alpha\beta} + 4\pi \frac{\partial \mathcal{P}_\alpha}{\partial \mathcal{E}_\beta} \tag{53}$$

which is measured experimentally at frequencies high enough to be ignored by the phonons but low enough not to excite electrons into the conduction band. As the first derivative of the total energy with respect to $\mathcal{E}$ is just $\mathcal{P}$, we can relate ε^{∞} to the second derivative of E with respect to $\mathcal{E}$, which is also the dielectric susceptibility χ.
The photon has no mass, such that its (relativistic) momentum is almost negligible compared to that of the phonons. Short of going to hard X rays or gamma radiation, we can consider that $\Delta\mathbf{q}$ due to the photon is 0, and the macroscopic $\mathcal{E}$ field considered here will be valid in general. The long wavelength limit is carried out [15] with a second Taylor expansion, with respect to the wavevector $\mathbf{q}$. The first non-zero term ($d\psi^{(1)}/dq_\alpha$) must be orthonormalized to $\psi^{(0)}$ but also to $d\psi^{(0)}/dk_\alpha$. This derivative with respect to wavevector (or momentum) should not surprise us: the position operator $\hat{r}$ (present in the electric field term of the Hamiltonian) can be linked by Fourier transform to the operator $d/d\mathbf{k}$ operating on a plane wave. The wave-vector derivatives $d\psi^{(0)}/dk_\alpha$ must be pre-calculated, but at little extra cost: the only terms with a contribution (those which do not commute with $d/d\mathbf{k}$) are the kinetic energy and an eventual non-local pseudopotential. In reciprocal space the derivatives of these terms are straightforward. The external potential is local, and in DFT the xc potential is as well: locality in real space implies no dependency on $\mathbf{k}$ and $d/d\mathbf{k} = 0$.
Once the $\psi^{\mathcal{E}_\alpha}$ have been calculated through a stationary expression like Eq. 33, the full matrix of second derivatives with respect to macroscopic electric field can be calculated from a simple non-stationary expression:

$$E^{\mathcal{E}_\alpha \mathcal{E}_\beta} = \frac{\Omega_0}{8\pi^3} \int_{\mathrm{BZ}} d\mathbf{k} \sum_j \langle i\psi^{(1)k_\alpha}_{\mathbf{k}j} | \psi^{(1)\mathcal{E}_\alpha}_{\mathbf{k}j} \rangle \tag{54}$$

where the superscripts on the ψ indicate which derivative/perturbation is being taken.

Born effective charges

The Born effective charge is defined as the change in polarization when an atom is displaced, or equivalently the change in force induced on a given atom when an electric field is applied:

$$Z^{\star}_{\kappa\beta\alpha} = \Omega_0 \frac{\partial \mathcal{P}_\beta}{\partial \tau_{\kappa\alpha}(\mathbf{q}=\mathbf{0})} = \frac{\partial F_{\kappa\alpha}}{\partial \mathcal{E}_\beta} \tag{55}$$

These values are tensorial since the induced charge may potentially give polarization changes in any direction β for an atomic displacement along α, depending on symmetry. For cubic and hexagonal systems the $Z^\star$ are usually diagonal. With the $\psi^{(1)}$ obtained for atomic-displacement perturbations, one can use a non-stationary expression to calculate $Z^\star$:

$$Z^{\star}_{\kappa\beta\alpha} = Z_\kappa \delta_{\beta\alpha} + 2\frac{\Omega_0}{8\pi^3}\int_{\mathrm{BZ}} d\mathbf{k} \sum_j \langle \psi^{(1)\tau_{\kappa\alpha}}_{\mathbf{k}j\mathbf{q}=\mathbf{0}} | i\psi^{(1)k_\beta}_{\mathbf{k}j}\rangle \tag{56}$$

where Z_κ is the (valence) ionic charge of atom κ.

Static dielectric response and LO/TO splitting

The Born effective charges are a univocal and quantitative way to assess charge associations in solids: if the second term in Eq. 56 is large something interesting is happening with bonding or electronic states in interaction with the phonons. The main use of $Z^\star$ is as an ingredient to describe the dielectric response in other quantities.

In the low frequency limit (photon energy comparable with the phonon ones) the dielectric response of solids includes the electronic part derived above but also the ionic part. We define $p_{m\gamma} = \sum_{\kappa\beta} Z^\star_{\kappa\gamma\beta}\eta_{m\mathbf{q}=\mathbf{0}}(\kappa\beta)$ to be the mode polarity, i.e. the nominal variation of the polarization along the phonon displacement vector. The full response is then:

$$\varepsilon^0_{\alpha\beta}(\omega) = \varepsilon^\infty_{\alpha\beta} + \frac{4\pi}{\Omega_0}\sum_m \frac{p_{m\alpha}p_{m\beta}}{\omega^2_{m\mathbf{q}=\mathbf{0}} - \omega^2} = \varepsilon^\infty_{\alpha\beta} + \frac{4\pi}{\Omega_0}\sum_m \frac{S_{m\alpha\beta}}{\omega^2_{m\mathbf{q}=\mathbf{0}} - \omega^2} \tag{57}$$

where S_m is the mode-dependent oscillator strength. This also provides the infrared absorption intensity

$$I(\omega) = \sum_m \delta(\omega-\omega_m)\mathrm{Tr}[S_m] = \sum_m \delta(\omega-\omega_m)\left|\sum_{\kappa\beta} p_{m\beta}\right|^2 \tag{58}$$

for an α polarized $\mathcal{E}$ field.

The infrared reflectivity follows directly from the determination of ε^0. One needs to project onto the direction of the (longitudinal) $\mathcal{E}$ field for an EM wave normal to a surface

$$\varepsilon_{\hat{\mathbf{q}}}(\omega) = \sum_{\alpha\beta} \hat{q}_\alpha \varepsilon^0_{\alpha\beta}(\omega)\hat{q}_\beta \tag{59}$$

to yield the reflectivity

$$R(\omega) = \left|\frac{\varepsilon^{1/2}_{\hat{\mathbf{q}}}(\omega) - 1}{\varepsilon^{1/2}_{\hat{\mathbf{q}}}(\omega) + 1}\right|^2 \tag{60}$$

One final consequence of the dynamical behavior of charges around atoms is that the frequency of certain long wavelength ($\mathbf{q} \to 0$) phonons will be affected in ionic compounds: if atomic displacements are longitudinal (i.e. along the direction of $\mathbf{q}$), a macroscopic electrical polarization may develop. This will cost energy, pushing the phonon frequency up. The correction term to the dynamical matrix (resp. frequencies) is proportional to the Born effective charges, and is screened by the electronic dielectric permittivity:

$$\begin{aligned} \tilde{\Phi}^{(2)}(\kappa\alpha;\kappa'\beta;\mathbf{q}\to 0) &= \tilde{\Phi}^{(2)}(\kappa\alpha;\kappa'\beta;\mathbf{q}=0) + \frac{4\pi}{\Omega_0}\frac{(\sum_\gamma q_\gamma Z^\star_{\kappa\gamma\alpha})(\sum_\delta q_\delta Z^\star_{\kappa'\delta\beta})}{\sum_{\gamma\delta} q_\gamma \varepsilon^\infty_{\gamma\delta} q_\delta} \\ \omega_m^{\mathrm{LO}} &= \omega_m^{\mathrm{TO}} + \frac{4\pi}{\Omega_0}\frac{(\mathbf{q}\cdot\mathbf{p}_m)^2}{\sum_{\gamma\delta} q_\gamma \varepsilon^\infty_{\gamma\delta} q_\delta} \end{aligned} \tag{61}$$

For acoustic modes all atoms move together, whether along $\mathbf{q}$ or not, and no polarization develops in any event. The optical modes must be separated into two sets. For a given $\mathbf{q}$ there will be two phonons which are transverse (labeled TO if $\mathbf{Z}^\star \cdot \mathbf{q} = 0$), where nothing happens, and one mode which is longitudinal (LO). For ionic systems there will be a splitting between the LO and TO frequencies at Γ.

Care must be taken with limits coming from different directions in $\mathbf{q}$. Due to the tensorial nature of $Z^\star$ the correction term is not analytic, and its value *at* $\mathbf{q} = 0$ will depend on the direction $\hat{\mathbf{q}}$ (note the norm of $\mathbf{q}$ simplifies in Eq. 61). This strange behavior is physical, and several sets of frequencies can be observed, e.g. by infrared spectroscopy, depending on the orientation of a crystal. The frequency of the observed phonons depends on the polarizability in the different directions, with respect to the propagation of the incoming electric field.

4 Other perturbations and developments

This section details a selection of extensions of DFPT to other perturbations: other external fields or other internal parameters of the DFT Hamiltonian. We also mention a few developments of the "normal" theory to include interesting electronic effects, in particular relativistic and spin effects for heavy and magnetic elements.

4.1 3rd order responses

Tensors of Raman intensities

Raman spectroscopy observes the shift in frequency of a laser passing through a material, usually in the infra-red or visible range [20]. As a chemical and structural analysis tool Raman is very widely used to characterize molecules, additives, biological complexes, minerals and many other materials, both experimentally and industrially. The shift in light frequency is due to the absorption or emission of a Γ-point phonon in the material (usually the only ones able to interact with photons), such that the distribution of Raman frequencies can be calculated from the harmonic phonon frequencies developed in the previous sections, choosing the irreducible representations which are active for the Raman interaction Hamiltonian. The symmetry of the specific phonon mode involved can be determined experimentally by polarizing the light coming in and out of the sample, giving selection rules for the full 3D tensor of Raman intensities $\aleph_{\alpha\beta}$ (see chapter 8 of Ref. [21]). The intensity of the absorption, however, is a more complex quantity, which includes the number of available phonon states (DOS) and the strength of the transition matrix element. Because of the quantum mechanics of harmonic oscillator creation and annihilation operators, the absorption is always favored (the shifts in frequencies are also known as Stokes and anti-Stokes). The transitions can be seen electromagnetically as a modulation of the dielectric constant of the medium by the phonons. We have detailed above that the (static) dielectric constant $\varepsilon_{\alpha\beta}$ can be calculated from the electronic response corrected by the ionic response given as a function of $Z^\star$, the Born effective charges (i.e. harmonic or second order derivatives). For the Raman intensities we need the variation of these quantities with respect to a phonon mode, and is therefore a 3rd order derivative of the energy. Luckily, using the 2n+1 theorem, we can extract this derivative from the 1st order perturbed wavefunctions, which are already provided by the normal harmonic phonon calculation [22].
The dielectric constant can be obtained as the second order derivative of the total energy with respect to macroscopic external electric fields: the first derivative gives the polarization, and the linear term in the polarization is proportional to $\varepsilon_{\alpha\beta}$, or equivalently to the polarizability $\chi_{\alpha\beta}$. The Raman scattering cross section per unit solid angle is given by

$$\frac{dS}{d\Omega} = (n_m + 1)\frac{h}{2\omega_m}\frac{(\omega_0 - \omega_m)^4}{c^4}|\mathbf{e}_s \cdot \aleph^m \cdot \mathbf{e}_0|^2 \tag{62}$$

where ω_m is the phonon energy, ω_0 the photon energy, $\mathbf{e}_s$ and $\mathbf{e}_0$ the outgoing and incoming polarizations, and n_m is the Bose-Einstein occupation factor for the phonon under consideration. The central DFPT quantity is the third order derivative of the energy with respect to two electric fields $\mathcal{E}$ (i.e. the first order dielectric susceptibility) and one atomic position $\tau_{\kappa\beta}$ (atom κ and direction β).

$$\frac{\partial \chi^{\infty(1)}_{\alpha\beta}}{\partial \tau_{\kappa\beta}} = \left.\frac{\partial^3 E}{\partial \mathcal{E}_i \partial \mathcal{E}_j \partial \tau_{\kappa\beta}}\right|_{\mathcal{E}=0} - \frac{8\pi}{\Omega_0}\frac{\sum_l Z^\star_{\kappa\beta\gamma} q_\gamma}{\sum_{\gamma\gamma'} q_\gamma \varepsilon^\infty_{\gamma\gamma'} q_{\gamma'}} \sum_\gamma \chi^{\infty(2)}_{\alpha\beta\gamma} q_\gamma \tag{63}$$

where we have added the non-analytic contribution for polar crystals (the most common case) in which $Z^\star$ is the Born effective charge, $\mathbf{q}$ the wavevector of the photon, which is only important by its direction, ε^∞ is the optical dielectric constant as above, and $\chi^{\infty(2)}_{\alpha\beta\gamma}$ is the non-linear optical susceptibility.
The Raman tensor is then given by

$$\aleph^m_{\alpha\beta} = \sqrt{\Omega_0} \sum_{\kappa\gamma} \frac{\partial \chi^{\infty(1)}_{\alpha\beta}}{\partial \tau_{\kappa\gamma}} \eta_m(\kappa\gamma) \tag{64}$$

where $\eta_m(\kappa\gamma)$ is the displacement vector for phonon mode m. In powder, polycrystalline, or disordered samples appropriate averages of $\aleph$ must be taken to compare to experiment. The theory exposed so far has described non-resonant Raman, and presumed no electronic transitions are available at the same photon energies. In many materials, in particular metals, this is not true, as there is no energy gap in the electron spectrum. The generalization to resonant Raman is an active field of research, e.g. for DFPT in a recent paper by Gillet *et al.* [23].
Note finally that $\aleph$ gives the intensity of the Raman peaks (in particular it tells you if for some Raman-allowed modes $\aleph^m$ is small or 0) but does not give the broadening or line width. To this end one must explicitly consider the processes which limit phonon lifetimes, in particular the electron-phonon and phonon-phonon interactions, which will be described below.

Electro-optic coefficients

A closely related quantity is the elecro-optic coefficient, which is the variation of dielectric constant ε under applied external electric field. The field will give a direct effect through $\chi^{\infty(2)}$ (the third partial derivative with respect to electric field) and an indirect effect when the atomic positions relax under the field. The latter can be estimated to first order from the phonon and electric field derivatives we already have. The variation of ε

$$\Delta(\varepsilon^{-1})_{\alpha\beta} = \sum_{\gamma=1}^{3} r_{\alpha\beta\gamma} \mathcal{E}_\gamma \tag{65}$$

is given by the Pockels coefficient $r = r^{el} + r^{ion}$

$$r^{el}_{\alpha\beta\gamma} = \frac{-8\pi}{\mathbf{n}^2_\alpha \mathbf{n}^2_\beta} \chi^{\infty(2)}_{\alpha\beta\gamma} \tag{66}$$

$$r^{ion}_{\alpha\beta\gamma} = -\frac{4\pi}{\sqrt{\Omega_0}\mathbf{n}^2_\alpha \mathbf{n}^2_\beta} \sum_m \frac{\aleph^m_{\alpha\beta} p_{m\gamma}}{\omega^2_m} \tag{67}$$

where $\mathbf{n}$ is the refractive index and $p_{m\gamma}$ is the mode polarity as in section 3.3.
If, further, the strain changes under applied field (which is generically always the case) a full set of linear equations must be solved for the so-called "unclamped" effect of the external field, including piezoelectric and strain perturbation effects. For finite electric fields within the Berry phase formulation of the modern theory of polarization [19] (see Lecture A6) it is not trivial whether to first discretize the calculation of the Berry phase then do the perturbation theory or the reverse. Veithen et al. [22] have shown the former (Perturbation Expansion After Discretization) is more efficient in terms of convergence, though both give the same numerical results.

Anharmonic inter-atomic force constants

The lattice dynamics of real crystals is never fully harmonic. Interatomic potential energies and the corresponding force constants are always anharmonic beyond a certain distance of stretching, and often beyond just a small fraction of an interatomic distance. The anharmonicity of these potentials leads to lower forces upon stretching bonds, and when a crystal is heated and more phonons are excited the distances expands. With respect to the ideal harmonic phonon picture, the presence of a non-harmonic potential gives rise to (quasi)-particles with a finite lifetime. Spectroscopic features (e.g. Raman) due to a phonon will no longer be infinitely sharp, but have a width in frequency/energy, which is linked to the inverse of the lifetime through a Heisenberg relation.

Thermal expansion is the first visible consequence of anharmonicity, but there are many others. The thermal conductivity of solids is determined mainly by phonon heat transfer, limited by phonon-phonon interactions (the "lattice" contribution). This is true both in insulators and, for intermediate temperatures, in metals - at very high T the electronic contribution dominates as there are often many more degrees of freedom in the electrons. The thermal variation of any experimental observable is a combination of a term due to thermal expansion, and additional terms due to direct interaction with phonons. Thermal motion guarantees that some of the anharmonic energy landscape will be sampled by atomic trajectories, and in many cases can dynamically stabilize or destabilize a crystal. One such case is Calcium under pressure, where a simple cubic phase is stabilized by anharmonic renormalizations of the phonon frequencies: a tetragonal phase with lower internal energy is rendered unstable at finite temperature [24] with full anharmonic renormalization (NB this is beyond DFPT), and may even be less stable in internal energy at low temperature, if full quantum effects are taken into account [25] (NB this builds on developments in the present section).

The full energy of a crystal can be written in a Taylor series (as Eq. 36 above) but with an extra term written explicitly),

$$E = E_0 + \sum_{\mathbf{R}\kappa\beta} \Phi^{(1)}(\mathbf{R}\kappa\beta) u_{\mathbf{R}\kappa\beta} + \frac{1}{2} \sum_{\mathbf{R}\kappa\beta\mathbf{R}'\kappa'\beta'} \Phi^{(2)}(\mathbf{R}\kappa\beta;\mathbf{R}'\kappa'\beta') u_{\mathbf{R}\kappa\beta} u_{\mathbf{R}'\kappa'\beta'} + \frac{1}{3!} \sum_{\mathbf{R}\kappa\beta\mathbf{R}'\kappa'\beta'\mathbf{R}''\kappa''\beta''} \Phi^{(3)}(\mathbf{R}\kappa\beta;\mathbf{R}'\kappa'\beta';\mathbf{R}''\kappa''\beta'') u_{\mathbf{R}\kappa\beta} u_{\mathbf{R}'\kappa'\beta'} u_{\mathbf{R}''\kappa''\beta''} + \ldots \quad (68)$$

with the various orders of force constants Φ and displacements u of the atomic positions τ in the unit cell at $\mathbf{R}$, and κ runs over all atoms in the cell. If atoms start from their equilibrium positions the linear term (the force) is 0 and the second order (harmonic) force constants are those calculated in Section 3.1. For temperatures well below melting, the u can be considered small parameters and the series converges, each higher order term being smaller than the previous one. Near melting, or in ionic conductors, atoms can diffuse and the u are by no means small, or even bounded, and the series expansion is not well defined. With these limitations, DFPT can in principle calculate all of the $\Phi^{(n)}$, following the ladder of variational approximations and 2n+1 theorems exposed above, but the increasing complexity of the formulae and the associated phase space (number of $\mathbf{q}$ vectors) limit calculations to $\Phi^{(3)}$ for the moment - these are also the most important contributions in almost all cases. Combinations of DFPT with additional finite differentiation have yielded $\Phi^{(4)}$ for some simple systems, such as in Ref. [25].

The theory for the calculation of $\Phi^{(3)}$ was implemented by Alberto Debernardi [26, 27] for

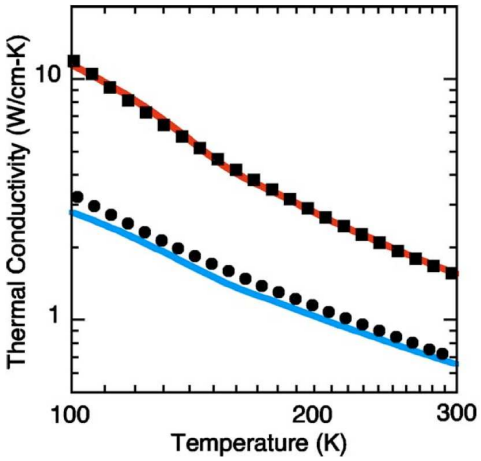


Fig. 1: *Thermal conductivity of Silicon (top red curve) and Germanium (bottom blue curve), compared to experimental data. From Reference [30]. Excellent agreement is obtained with 3rd order phonon-phonon scattering, plus isotopic and defect contributions (both of which have much simpler forms).*

Γ-point phonons, yielding directly the derivative

$$\Phi^{(3)}(\mathbf{R}\kappa\beta;\mathbf{R}'\kappa'\beta';\mathbf{R}''\kappa''\beta'') = \frac{\partial^3 E}{\partial\tau_{\mathbf{R}\kappa\beta}\partial\tau_{\mathbf{R}'\kappa'\beta'}\partial\tau_{\mathbf{R}''\kappa''\beta''}} \tag{69}$$

(each atom κ is in unit cell $\mathbf{R}$ and displaced in direction β) or more often its lattice Fourier transform:

$$\Phi^{(3)}_{\kappa\beta\kappa'\beta'\kappa''\beta''}(\mathbf{q}\mathbf{q}'\mathbf{q}'') = \frac{1}{N}\sum_{R=0R'R''} e^{i\mathbf{q}\cdot\mathbf{R}}e^{i\mathbf{q}'\cdot\mathbf{R}'}e^{i\mathbf{q}''\cdot\mathbf{R}''}\Phi^{(3)}(\mathbf{0}\kappa\beta;\mathbf{R}'\kappa'\beta';\mathbf{R}''\kappa''\beta'') \tag{70}$$

with the variation of κ restricted to a single unit cell at the origin and the use of translational invariance. Note that this partial FT with respect to lattice vectors is very common, and yields a hybrid quantity in reciprocal (for $\mathbf{q}$) and real space (for $\tau_{\mathbf{0}\kappa}$). Both energy and momentum must be conserved in the 3 phonon interaction:

$$\mathbf{q} + \mathbf{q}' + \mathbf{q}'' = \mathbf{G} \tag{71}$$

$$\omega_{\mathbf{q}m} + \omega_{\mathbf{q}'m'} \pm \omega_{\mathbf{q}''m''} = 0 \tag{72}$$

where the sign determines phonon emission or absorption. For Γ point phonons a significant reduction of the phase space happens, as $\mathbf{q}' = -\mathbf{q}''$ to within a $\mathbf{G}$ vector. This is the useful case for the Raman modes mentioned above. Extensions of the implementation to finite $\mathbf{q}$ were carried out in the Regensburg group by Deinzer, Strauch et al. [28, 29].
From the $\Phi^{(3)}$ one can calculate the phonon linewidth as a function of temperature (here at Γ for simplicity):

$$\begin{aligned}\Gamma_{m\mathbf{q}=0} &= \frac{\pi}{2\hbar^2}\sum_{\mathbf{q}'m'm''}\left|\sum_{\kappa\beta}\Phi^{(3)}_{\kappa\beta\kappa'\beta'\kappa''\beta''}(\mathbf{0}\ \ \mathbf{q}'-\mathbf{q}')\eta_{m0}(\kappa\beta)\eta_{m'\mathbf{q}'}(\kappa'\beta')\eta_{m''\mathbf{q}''}(\kappa''\beta'')\right|^2 \\ &\times \ [(1+n_{\mathbf{q}'m'}+n_{-\mathbf{q}'m''})\delta(\omega_{\mathbf{q}'m'}+\omega_{-\mathbf{q}'m''}-\omega_m) \\ &+ \ 2(n_{\mathbf{q}'m'}-n_{-\mathbf{q}'m''})\delta(\omega_{\mathbf{q}'m'}-\omega_{-\mathbf{q}'m''}-\omega_m)]\end{aligned} \tag{73}$$

and other quantities such as Grüneisen parameters [31] through similar formulae:

$$\gamma_{m\mathbf{q}} = -\frac{1}{6\omega^2_{m\mathbf{q}}} \sum_{\kappa\beta;\mathbf{R}';\kappa'\beta'\mathbf{R}''\kappa''\beta''} \Phi^{(3)}(\mathbf{0}\kappa\beta;\mathbf{R}'\kappa'\beta';\mathbf{R}''\kappa''\beta'')\eta^*_{m\mathbf{q}}(\kappa\beta)\eta_{m\mathbf{q}}(\kappa'\beta')e^{i\mathbf{q}\cdot\mathbf{R}'}\tau_{\mathbf{R}''\kappa''\beta''}$$

The positions τ in the final factor will be renormalized by the thermal phonon displacements if the symmetry is low enough - see Ref. [31]. Finally, the thermal conductivity was first determined from DFPT by Broido et al. in Ref. [30] - see Fig. 1. In many systems $\Phi^{(3)}$ gives the bulk of the anharmonicity and very good agreement can be obtained, once isotopic and impurity scattering is added to the phonon contribution.

4.2 Spin and magnetic effects

Magnetic field and magnons

The second variation of the energy with respect to an external magnetic field gives the magnetic susceptibility χ_m. In non magnetic materials, paradoxically, the (dia)magnetic response is more complex to calculate. This response determines the nuclear magnetic resonance (NMR) or electron paramagnetic resonance (EPR) shifts, and is the subject of ongoing research [32, 33, 34]. In ferromagnetic or antiferromagnetic systems (in particular metals, but sometimes insulators as well) the χ_m presents poles for excitations of the magnetic system (waves of oscillating spins or magnons). These excitations are quite low in energy (meV or much less). Because the degree of freedom (the spin) being perturbed is intimately linked to the electron wave functions, one can not use a phonon-like separation of ionic and electronic coordinates to study "just" the effect of the $\mathbf{B}$ field perturbatively. The correct formulation was developed by Savrasov [35] as a perturbation theory built on top of time dependent DFT [36] (see Lecture A7). The action is varied perturbatively, instead of the total energy, under an applied *time dependent* magnetic field

$$\delta B_{\text{ext}}(\mathbf{r}t) = \delta\mathbf{b}e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}e^{i\omega t} \tag{74}$$

The first order change in wave functions induces changes in density *and* magnetization:

$$\delta\mathbf{m} \;=\; \mu_B \sum_{iss'} \psi^{(1)}_{is}\sigma_{ss'}\psi^{(0)}_{is'} + \text{c.c.} \tag{75}$$

summing over bands i and spins s, where μ_B is the Bohr magneton and σ are the Pauli matrices. The change $\delta\mathbf{m}$ for a given field yields χ_m. The 2n+1 and variational principles are extended, giving the time dependent Sternheimer equation, conveniently written in frequency space:

$$(H - \epsilon_i \pm \omega)\psi^{(1)}_i + (V^{(1)}_{eff}I - \mu_B\sigma\delta B_{xc})\psi^{(1)}_i = 0 \tag{76}$$

where I is the 2×2 spin identity matrix. There is a further feedback effect through the exchange-correlation magnetic field δB_{xc}, which is induced by the DFT electrons themselves.
Magnon bands are seen as peaks or divergencies of χ_m at specific energies, but $\chi_m(\omega)$ is a continuous function giving the full response at all frequencies. The equations must be solved for the full $\mathbf{q}\,\omega$ space, which makes things quite heavy numerically. Few systems with more than 1 atom per unit cell have been calculated in this way. Other approaches such as the Korringa Kohn Rostoker [37] Green's function approach are more widely used for magnon dispersions.

Spin-orbit coupling

In heavy atoms the Coulomb potential is extremely strong near the nucleus. Core electrons see a very large electric field, and feel a very strong potential. This quantum confinement implies a correspondingly high kinetic energy, and core electrons can quickly attain relativistic speeds, obeying the Dirac 4-component spinor equation instead of the "classical" Schrödinger one (see Ref. [38] and Lecture A10). In the frame of reference of the electrons, the electric field is equivalent to a magnetic one, which affects the spin angular momentum of the electron. Depending on the orbital angular momentum L of the electrons, they will have an angular speed, and a Taylor expansion of the Dirac equation leads to scalar corrections and then the famous spin-orbit coupling (SOC) term, proportional to $\mathrm{L}\cdot\mathrm{s}$ with a prefactor proportional to ∇v_{ext} (i.e. the electric field).
The coupling between spin and angular orbital moments is almost always treated in DFT through a reduced equation for 2-component Pauli spinors [39] with spin up and spin down components, instead of the full Dirac equation. In this case the scalar and spin-orbit relativistic terms are simply additive in the normal Kohn-Sham Hamiltonian, and the extension of DFPT is quite straightforward. Only the spin-orbit term is not spin-component diagonal. In norm-conserving pseudopotential [40] and projector augmented wave [41] formalisms the effect on the active valence electrons is treated in a perturbative way: the valence electron wave functions are projected on an atomic basis, in which the matrix element of the SOC is evaluated. The effect can be wrapped into non-local pseudopotential like projectors, in a sphere around each atom. In this way, for phonon perturbations, the dependency on atomic positions is simple: the projector is rigidly shifted with the atom. In reciprocal space the phase factor corresponding to the non-local projection operators are simple to derive with respect to ionic positions. We consider a SOC pseudopotential term of the form:

$$V_l^{SO}(\mathbf{G}s,\mathbf{G}'s') = -i\frac{4\pi}{\Omega_0}(2l+1)\sum_{\kappa} E_{\kappa l}^{KB,SO} P_l'(\hat{G}\cdot\hat{G}')\zeta_{\kappa l}^{SO}(\mathbf{G})\zeta_{\kappa l}^{SO}(\mathbf{G}') \left(\langle s|\sigma|s'\rangle\cdot\frac{\mathbf{G}\times\mathbf{G}'}{GG'}\right) e^{i(\mathbf{G}'-\mathbf{G})\tau_\kappa} \tag{77}$$

in reciprocal space $\mathbf{G}$, where $\zeta_{\kappa l}^{SO}(\mathbf{G})$ is the spherical Fourier transform of the radial part of the potential, and with Kleinman Bylander energies E_l^{KB} and Legendre polynomials P_l for each angular momentum channel l. It becomes clear that the atomic position dependency is simple. The second derivative of the energy term becomes:

$$\frac{\partial^2 E^{SO}}{\partial\tau_{\kappa\alpha}\partial\tau_{\kappa\alpha}} = i\frac{4\pi}{\Omega_0}\sum_{nl} f_n(2l+1)E_{\kappa l}^{KB,SO}\sum_{\mathbf{G}s\mathbf{G}'s'} P_l'(\hat{G}\cdot\hat{G}')\zeta_{\kappa l}^{SO}(\mathbf{G})\zeta_{\kappa l}^{SO}(\mathbf{G}') \left(\langle s|\sigma|s'\rangle\cdot\frac{\mathbf{G}\times\mathbf{G}'}{GG'}\right)\times 2\mathfrak{Re}[(G'_\alpha-G_\alpha)G'_\beta\cdot c_{ns}^{(0)*}(\mathbf{G})c_{ns'}^{(0)}(\mathbf{G}')\cdot e^{i(\mathbf{G}'-\mathbf{G})\tau_\kappa}] \tag{78}$$

where f_n is the Fermi Dirac occupation and $c_{ns}(\mathbf{G})$ are the plane wave coefficients for $\psi_n^{(0)}$. The formalism has been developed for norm-conserving pseudopotentials in Verstraete *et al.* [42] and for PAW by Dal Corso [43]. For Pb the effect of spin orbit coupling on the phonons is particularly strong at the X point. The SOC produces a shift of the Fermi surface, part of which moves to half of the distance between X and the zone center. The resulting nesting emphasizes a Kohn anomaly [44] at X and softens the phonon frequency - see Fig. 2.

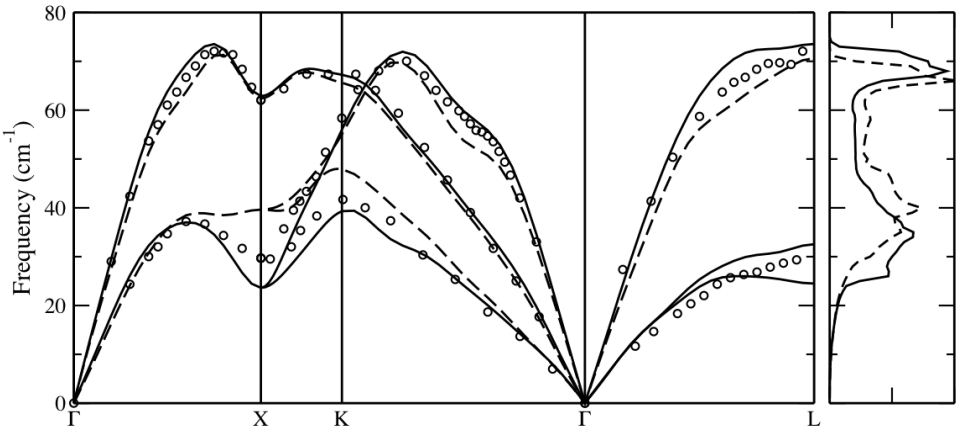


Fig. 2: *Phonon band structure and DOS of lead, with (solid) and without (dashed) spin-orbit coupling, compared to experiment (symbols). A huge softening at X results from changes in the Fermi surface due to SOC and a resulting Kohn anomaly. From Reference [42]*

4.3 Elastic constants

Another important physical response function is the elastic constant tensor, the second derivatives of the total energy (or rather the Helmholtz free energy) with respect to applied strain ς. As with all second derivatives they can be calculated from the first order perturbed wave functions, but the derivatives are much more involved for reasons we will see. A DFPT formalism was first written down by Baroni, Giannozzi and Testa [45] for isotropic stress and local pseudopotentials, and generalized more recently by Hamann *et al.* [46] (we will follow the latter). Because the strain can change the unit cell angles as well as the cell parameters, the basis set itself will be affected, and all of the terms of the energy will vary. The formalism is expressed as a function of the reduced atomic coordinates and the metric tensors:

$$\Xi_{ij} = \sum_{\alpha} R_{\alpha i} R_{\alpha j} \tag{79}$$

$$\Upsilon_{ij} = \sum_{\alpha} G_{\alpha i} G_{\alpha j} \tag{80}$$

in real and reciprocal space, with α here a cartesian coordinate, i j reduced coordinates (we use roman letters to denote reduced coordinates in this section for clarity), and R and G the real and reciprocal space lattice vectors. The main advantage is that in reduced coordinates the boundary conditions do not change, which is not the case in cartesian coordinates: any finite strain produces arbitrarily large shifts for atoms far enough away from the origin. The derivatives of Υ (idem for Ξ) with respect to strain are apparently simple:

$$\Upsilon_{ij}^{(\alpha\beta)} = \frac{\partial \Upsilon_{ij}}{\partial \varsigma_{\alpha\beta}} = -G_{\alpha i} G_{\beta j} - G_{\beta i} G_{\alpha j} \tag{81}$$

$$\begin{aligned} \Upsilon_{ij}^{(\alpha\beta\gamma\delta)} = \frac{\partial^2 \Upsilon_{ij}}{\partial \varsigma_{\alpha\beta} \partial \varsigma_{\gamma\delta}} &= \delta_{\alpha\gamma}(G_{\beta i} G_{\delta j} + G_{\delta i} G_{\beta j}) + \delta_{\beta\gamma}(G_{\alpha i} G_{\delta j} + G_{\delta i} G_{\alpha j}) \\ &+ \delta_{\alpha\delta}(G_{\beta i} G_{\gamma j} + G_{\gamma i} G_{\beta j}) + \delta_{\beta\delta}(G_{\alpha i} G_{\gamma j} + G_{\gamma i} G_{\alpha j}) \end{aligned} \tag{82}$$

The dependency on Ξ and Υ can be made explicit in each of the energy terms, as well as the wave function normalization expressions. We will not go through all of the derivatives, but

show as an example the effect on the kinetic energy T, which is not directly affected by *any* other perturbation (only indirectly through variations of the wave functions). The contribution of individual plane waves to T is:

$$\langle\tilde{\mathbf{K}}'|T|\tilde{\mathbf{K}}\rangle = \frac{1}{2}\delta_{\tilde{\mathbf{K}}\tilde{\mathbf{K}}'}\sum_{ij}\Upsilon_{ij}\tilde{K}_i\tilde{K}_j \tag{83}$$

where the tilde denotes reduced vectors. The first derivatives (stress contributions) are then:

$$\langle\tilde{\mathbf{K}}'|\frac{\partial T}{\partial\varsigma_{\alpha\beta}}|\tilde{\mathbf{K}}\rangle = \frac{1}{2}\delta_{\tilde{\mathbf{K}}\tilde{\mathbf{K}}'}\sum_{ij}\Upsilon_{ij}^{(\alpha\beta)}\tilde{K}_i\tilde{K}_j \tag{84}$$

The second derivatives (elastic constant contributions) are:

$$\langle\tilde{\mathbf{K}}'|\frac{\partial^2 T}{\partial\varsigma_{\alpha\beta}\partial\varsigma_{\gamma\delta}}|\tilde{\mathbf{K}}\rangle = \frac{1}{2}\delta_{\tilde{\mathbf{K}}\tilde{\mathbf{K}}'}\sum_{ij}\Upsilon_{ij}^{(\alpha\beta\gamma\delta)}\tilde{K}_i\tilde{K}_j \tag{85}$$

As can be seen the diagonal character of the kinetic energy and its simple form are preserved in reduced coordinates. Corresponding terms can be written for the Hartree, exchange-correlation, and pseudopotential energies, and a variational second order expression found which can be minimized to give $\psi^{(1)}$, then the elastic constants (supposing you start from an unstressed equilibrium state) are:

$$\bar{C}_{\alpha\beta;\gamma\delta} = \frac{1}{\Omega_0}\frac{\partial^2 E}{\partial\varsigma_{\alpha\beta}\partial\varsigma_{\gamma\delta}} \tag{86}$$

Combining the strain derivatives with electric field derivatives, one can access the piezoelectric coefficients (specifically the frozen ion piezoelectric tensor) through the mixed second derivative:

$$\bar{e}_{\gamma;\alpha\beta} = -\frac{\partial^2 E}{\partial\tilde{\mathcal{E}}_\gamma\partial\varsigma_{\alpha\beta}} = -2\frac{\Omega_0}{(2\pi)^3}\int_{BZ}\sum_m^{occ}\langle\psi_{\mathbf{k}m}^{(\tilde{k}_\gamma)}|\psi_{\mathbf{k}m}^{(\varsigma_{\alpha\beta})}\rangle d\mathbf{k} \tag{87}$$

in which the first perturbed wave function is derived with respect to the k-point coordinates (reduced), as above for other $\mathcal{E}$ field derivatives, and the second with respect to strain.
Under applied strain all of the cartesian atomic positions will change as well. The reduced coordinates can often stay the same (e.g. a centered atom in a BCC unit cell under isotropic dilation), but, even if the system is very highly symmetric, a shear strain will usually break enough symmetries to allow additional internal degrees of freedom. In this case one must couple the strain derivatives to phonon derivatives to calculate the *relaxed ion* elastic tensor. The systematic treatment of phonon, electric field, and strain perturbation combinations is given in Ref. [47]. Quantities calculated with frozen ions are denoted with a bar. The internal strain tensor is defined as

$$\Lambda_{m;\alpha\beta} = -\Omega_0\sum_{\kappa\gamma}\eta_{\mathbf{q}=\mathbf{0}m}(\kappa\gamma)\frac{\partial^2 E}{\partial\tau_{\kappa\gamma}\partial\varsigma_{\alpha\beta}} \tag{88}$$

and the resulting relaxed ion elastic and piezoelectric tensors are:

$$C_{\alpha\beta;\gamma\delta} = \bar{C}_{\alpha\beta;\gamma\delta} - \Omega_0^{-1}\Lambda_{m;\alpha\beta}((\Phi^{(2)})^{-1})_{mm'}\Lambda_{m';\gamma\delta} \tag{89}$$

$$e_{\gamma;\alpha\beta} = \bar{e}_{\gamma;\alpha\beta} + \Omega_0^{-1}Z^\star_{m\gamma}((\Phi^{(2)})^{-1})_{mm'}\Lambda_{m';\alpha\beta} \tag{90}$$

using the Born effective charges and the inverse of the harmonic force constants.

4.4 Electron-Phonon coupling

Framework theory

The response functions described up to this point determine the reaction of a system to perturbations, either of electrons or ions to external fields, or of electrons to the movement of the ions. We have supposed that the independent particle Kohn Sham picture is a good starting point, and that the corresponding single particle orbitals are good representations of the electronic states and excitations. Similarly we have supposed (except in Sect. 4.1) that the phonons are *also* ideal quasiparticles, with infinite lifetimes at a given volume and temperature. In reality, neither phonons nor electrons are ideal, and their main interaction is usually with each other - the presence of phonons (and many other scattering centers) limits the lifetime and free path of electrons, and vice versa. The electron phonon coupling (EPC) is central to "ordinary" BCS superconductivity, normal phase electron resistivity, and many other phenomena. The theory of EPC has been reviewed extensively by Allen and Mitrovic [48] and in a book by Grimvall [49]. The first DFT calculations in this framework were carried out by Marvin Cohen and coworkers in the 1980s [50]. The extension to DFPT is due to Savrasov [51]. Within the phonon perturbation theory exposed in the preceding sections we already have access to the necessary matrix element, which is

$$
\begin{aligned}
g^{\mathbf{q}m}_{\mathbf{k}+\mathbf{q}i'\mathbf{k}i} &= \langle \psi^{(0)}_{\mathbf{k}+\mathbf{q}i'} | \delta^{\mathbf{q}m} V_{eff} | \psi^{(0)}_{\mathbf{k}i} \rangle && (91)\\
&= \frac{1}{\sqrt{\omega_{\mathbf{q}m}}} \sum_{\kappa\alpha} \eta_{\mathbf{q}m}(\kappa\alpha) \frac{\partial V_{eff}}{\partial \tau_{\kappa\alpha}} && (92)\\
&= \frac{1}{\sqrt{\omega_{\mathbf{q}m}}} \sum_{\kappa\alpha} \eta_{\mathbf{q}m}(\kappa\alpha) \langle \psi^{(0)}_{\mathbf{k}+\mathbf{q}i'} | H^{(1)}_{\mathbf{q}\kappa\alpha} | \psi^{(0)}_{\mathbf{k}i} \rangle && (93)
\end{aligned}
$$

for phonon mode m at wave vector $\mathbf{q}$, and momentum conservation imposes $\mathbf{k}' = \mathbf{k} + \mathbf{q}$ (we will impose energy conservation later). The phonon displacement vector $\eta_{\mathbf{q}m}(\kappa\alpha)$ is in units of meters (i.e. it is normalized with the square roots of atomic masses). The matrix element is nothing but $\epsilon^{(1)\mathbf{q}\kappa\alpha}$ from the Sternheimer equation. It contains the self consistent induced change in the effective KS potential. Within the Migdal approximation the self-interaction terms for EPC are only considered with one phonon propagator [52], so this matrix element is all we need. Through Fermi's Golden Rule, the interaction will determine an inverse lifetime for phonons (by interaction with electrons with Fermi level ϵ_F):

$$
\gamma^{ph}_{\mathbf{q}m} = 2\pi\omega_{\mathbf{q}m} \sum_{\mathbf{k}ii'} |g^{\mathbf{q}m}_{\mathbf{k}+\mathbf{q}i'\mathbf{k}i}|^2 \delta(\epsilon_{\mathbf{k}i} - \epsilon_F)\delta(\epsilon_{\mathbf{k}+\mathbf{q}i'} - \epsilon_F) \qquad (94)
$$

and for electrons (by interactions with phonons - see e.g. Ref. [53, 49]):

$$
\begin{aligned}
\gamma^{el}_{\mathbf{k}i} &= 2\pi \sum_{\mathbf{q}mi'} |g^{\mathbf{q}m}_{i'\mathbf{k}+\mathbf{q}i\mathbf{k}}|^2 \times \{[f(\epsilon_{\mathbf{k}+\mathbf{q}i'}) + n(\omega_{\mathbf{q}m})]\delta(\epsilon_{\mathbf{k}i} - \epsilon_{\mathbf{k}+\mathbf{q}i'} + \omega_{\mathbf{q}m})\\
&+ [1 - f(\epsilon_{\mathbf{k}+\mathbf{q}i'}) + n(\omega_{\mathbf{q}m})]\delta(\epsilon_{\mathbf{k}i} - \epsilon_{\mathbf{k}+\mathbf{q}i'} + \omega_{\mathbf{q}m})\} \qquad (95)
\end{aligned}
$$

In the former we have made an additional approximation of elastic interaction, neglecting the phonon energy before the electron one, and restricting both initial and final states to be on the Fermi surface. In the latter one sees explicitly the terms for phonon emission and absorption, respectively.

Superconductivity

The full theory of phonon driven superconductivity is due to Eliashberg, who wrote a full microscopic and Green's function theory for the BCS interaction. The single term self-energy gives a spectral function analogous to a density of states, weighted by the EPC, known as the Eliashberg function:

$$\alpha^2F(\omega) = \frac{1}{2\pi N(\epsilon_F)} \sum_{\mathbf{q}m} \frac{\gamma_{\mathbf{q}m}}{\omega_{\mathbf{q}m}} \delta(\omega - \omega_{\mathbf{q}m}) \tag{96}$$

whose first inverse moment is the EPC coupling strength λ, which can be injected in your favorite formula for predicting superconducting critical temperatures (McMillan [54], Allen Dynes [55] ...).
The historical shortcoming of such formulae, and Eliashberg's final formulation as well, is the need for an empirical parameter characterizing the Coulomb repulsion between the Cooper-paired electrons in the system. The repulsion is well beyond the DFT interaction as it concerns two explicit electrons, and no mean field can provide a useful value. This problem has been addressed by the group of Hardy Gross [56, 57] using diagrammatic techniques on top of an explicit density functional theory for superconductors. Two density components are considered, for normal and superconducting electrons, and the coupling between the normal and "anomalous" densities are included. Very good agreement has been found in a number of systems (e.g. [58]).

Electron transport

The EPC matrix element also provides the main ingredient for theories studying electron transport. Phonon limitation of electrical conductivity can be included in any perturbational technique, but usually [59] starts from the semiclassical/mesoscopic Boltzmann transport equations (BTE). A number of seminal papers on EPC and BTE expressions were written by P.B. Allen since the 1970s [60, 61] and implemented in the work by Savrasov [51]. With a variational ansatz for the way the electron distribution is pushed out of equilibrium (away from Fermi Dirac) one can find a closed form for the transport coefficients. Allen introduces a generalized spectral function for transport, analogous to α^2F but including Fermi velocity factors:

$$\alpha^2_{\mathrm{tr}}F(\alpha,\beta,\omega) = \frac{1}{2N(\epsilon_F)} \sum_{m;\mathbf{k}i\mathbf{k}'i'} |g^{\mathbf{q}m}_{\mathbf{k}i\mathbf{k}'i'}|^2 \frac{[v_\alpha(\mathbf{k}) - v_\alpha(\mathbf{k}')]}{\langle v_\alpha \rangle} \frac{[v_\beta(\mathbf{k}) - v_\beta(\mathbf{k}')]}{\langle v_\beta \rangle} \delta(\epsilon_{\mathbf{k}i})\delta(\epsilon_{\mathbf{k}'i'})\delta(\omega_{\mathbf{q}m} - \omega)$$

where N is the density of states, and $\mathbf{v}(\mathbf{k})$ is the Fermi velocity at $\mathbf{k}$, $\langle v_\beta \rangle$ the FS average of $\mathbf{v}$ and ϵ_F has been set to 0. In the elastic approximation mentioned above, and to lowest order, the electrical resistivity is:

$$\rho_{\alpha\beta} = \frac{2\pi\Omega_0 k_B T}{N(\epsilon_F)\langle v_\alpha(\epsilon_F)\rangle\langle v_\beta(\epsilon_F)\rangle} \times \int_0^\infty \frac{d\omega}{\omega} \frac{x^2}{\sinh^2 x} \alpha^2_{\mathrm{tr}}F(\alpha,\beta,\omega) \tag{97}$$

where $x = \omega/2k_BT$. The electronic thermal resistivity (without the lattice contribution) has a similar form.

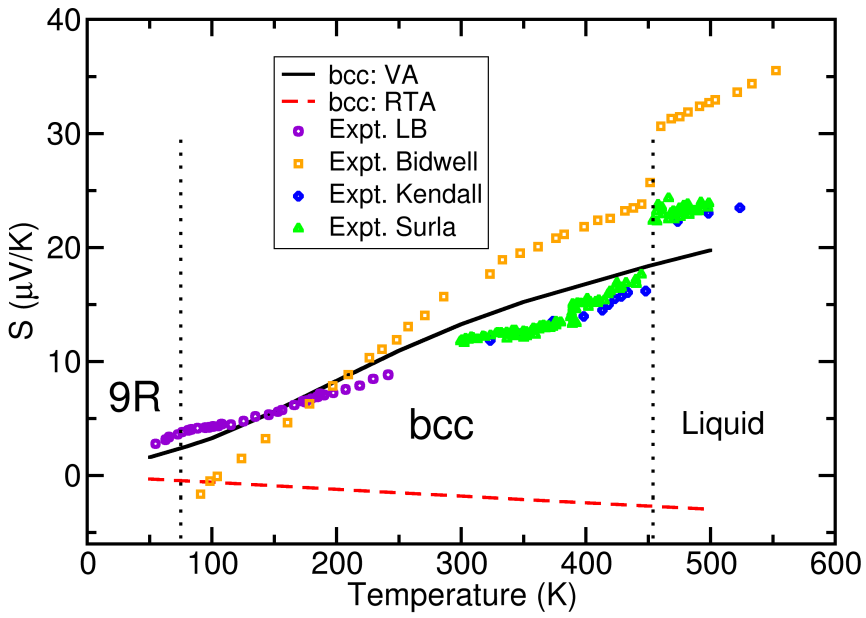


Fig. 3: *Seebeck coefficient of BCC lithium within DFPT (black) compared to experimental data (symbols) and constant relaxation time (red dashed). The values are positive, despite the negative sign of the charge carriers in this simple metal, due to a complex energy dependency of the electron lifetime. From Reference [62]*

Seebeck coefficients

Within the elastic approximation the Seebeck coefficient is identically zero, as electrons and holes will diffuse in exactly the same way at ϵ_F. Fermi smearing effects must be included, and the full expressions derived by Allen employed [61]. This was done recently [62] for BCC Li which presents an anomalous sign of the Seebeck coefficient. Standard approximations such as the constant relaxation time fail qualitatively (see Fig. 3).

Band gap renormalization

The interaction with phonons also has a more direct effect on the electronic energies, renormalizing them both through thermal expansion and through ionic screening and polarization effects. These are beyond the first order EPC described above, and were first codified by Allen, Heine and Cardona [63, 64] for empirical potentials. In the past few years several teams [65, 66] have implemented the DFPT calculation of the necessary terms and published results for diamond. There is some debate about the completeness of the AHC formalism and how close the DFT values should be to experiment.

References

[1] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **58**, 1861 (1987).

[2] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. Dal Corso, S. de Gironcoli, *et al.*, J Phys: Cond Matt **21**, 395502 (2009).

[3] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, *et al.*, Comp. Mat. Sci. **25**, 478 (2002).

[4] G. Kresse and J. Furthmller, Phys. Rev. B **54**, 11169 (1996).

[5] X. Andrade, J. Alberdi-Rodriguez, D. A. Strubbe, M. J. T. Oliveira, F. Nogueira, A. Castro, J. Muguerza, A. Arruabarrena, S. G. Louie, A. Aspuru-Guzik, A. Rubio, and M. A. L. Marques, J. Phys.: Cond. Matt. **24**, 233202 (2012).

[6] M. D. Segall, P. J. D. Lindan, M. J. Probert, C. J. Pickard, P. J. Hasnip, S. J. Clark, and M. C. Payne, J. Phys.: Cond. Matt **14**, 2717 (2002).

[7] X. Gonze, G.-M. Rignanese, and R. Caracas, Zeit. für Krist. **220**, 458 (2005).

[8] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).

[9] R. Caracas and X. Gonze, Zeit. für Krist. **220**, 511 (2005).

[10] H. Hellmann, *Einführung in die Quantumchemie* (Deuticke, Leipzig, Leipzig, 1937).

[11] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[12] R. M. Sternheimer, Phys. Rev. **96**, 951 (1954).

[13] X. Gonze and J.-P. Vigneron, Phys. Rev. B **39**, 13120 (1989).

[14] X. Gonze, D. C. Allan, and M. P. Teter, Phys. Rev. Lett. **68**, 3603 (1992).

[15] X. Gonze, Phys. Rev. B **55**, 10337 (1997).

[16] S. de Gironcoli, Phys. Rev. B **51**, 6773 (1995).

[17] M. Born and J. R. Oppenheimer, Ann. Phys. (Leipzig) **84**, 457 (1927).

[18] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).

[19] R. Resta, J Phys: Cond Matt **22**, 123201 (2010).

[20] N. B. Colthup, L. H. Daly, and S. E. Wiberley, *Introduction to Infrared and Raman Spectroscopy* (Academic Press, San Diego (CA), 1990).

[21] M. S. Dresselhaus, G. Dresselhaus, and A. Jorio, *Group theory* (Springer, Berlin ; Heidelberg, 2008).

[22] M. Veithen, X. Gonze, and P. Ghosez, Phys. Rev. B **71**, 125107 (2005).

[23] Y. Gillet, M. Giantomassi, and X. Gonze, Phys. Rev. B **88**, 094305 (2013).

[24] M. Di Gennaro, S. K. Saha, and M. J. Verstraete, Phys. Rev. Lett. **111**, 025503 (2013).

[25] I. Errea, B. Rousseau, and A. Bergara, Phys. Rev. Lett. **106**, 165501 (2011).

[26] A. Debernardi and S. Baroni, Sol. State. Comm. **91**, 813 (1994).

[27] A. Debernardi, *Anharmonic Properties of Semiconductors from Density-Functional Perturbation Theory*, Ph.D. thesis, SISSA, Trieste (1995).

[28] G. Deinzer, G. Birner, and D. Strauch, Phys. Rev. B **67**, 144304 (2003).

[29] G. Deinzer, M. Schmitt, A. P. Mayer, and D. Strauch, Phys. Rev. B **69**, 014304 (2004).

[30] D. A. Broido, M. Malorny, G. Birner, N. Mingo, and D. A. Stewart, App. Phys. Lett. **91**, 231922 (2007).

[31] J. Fabian and P. B. Allen, Phys. Rev. Lett. **79**, 1885 (1997).

[32] C. J. Pickard and F. Mauri, Phys. Rev. B **63**, 245101 (2001).

[33] T. Thonhauser, D. Ceresoli, D. Vanderbilt, and R. Resta, Phys. Rev. Lett. **95**, 137205 (2005).

[34] J. W. Zwanziger and M. Torrent, Appl. Magn. Reson. **33**, 447 (2008).

[35] S. Y. Savrasov, Phys. Rev. Lett. **81**, 2570 (1998).

[36] M. Marques, N. Maitra, F. Nogueira, E. Gross, and A. Rubio, eds., *Fundamentals of Time-Dependent Density Functional Theory*, vol. 837 of *Lecture Notes in Physics* (Springer, Berlin, 2012).

[37] H. Ebert, D. Kodderitzsch, and J. Minar, Rep. Prog. Phys. **74**, 096501 (2011).

[38] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley Publishing Company, Reading, MA, 1994), Revised ed.

[39] A. H. MacDonald and S. H. Vosko, J. Phys. C: Sol. State Phys. **12**, 2977 (1979).

[40] C. Hartwigsen, S. Goedecker, and J. Hutter, Phys. Rev. B **58**, 3641 (1998).

[41] A. Dal Corso, Phys. Rev. B **82**, 075116 (2010).

[42] M. J. Verstraete, M. Torrent, F. Jollet, G. Zérah, and X. Gonze, Phys. Rev. B **78**, 045119 (2008).

[43] A. Dal Corso, J Phys: Cond Matt **20**, 445202 (2008).

[44] W. Kohn, Phys. Rev. Lett. **2**, 393 (1959).

[45] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **59**, 2662 (1987).

[46] D. R. Hamann, X. Wu, K. M. Rabe, and D. Vanderbilt, Phys. Rev. B **71**, 035117 (2005).

[47] X. Wu, D. Vanderbilt, and D. R. Hamann, Phys. Rev. B **72**, 035105 (2005).

[48] H. Ehrenreich, F. Seitz, and D. Turnbull, eds., *Theory of Superconducting T_c*, vol. 37 of *Solid State Phys.* (Academic Press, New York, 1982).

[49] G. Grimvall, *The electron phonon interaction in metals* (North-Holland, Amsterdam, 1981).

[50] M. M. Dacorogna, M. L. Cohen, and P. K. Lam, Phys. Rev. Lett. **55**, 837 (1985).

[51] S. Y. Savrasov and D. Y. Savrasov, Phys. Rev. B **54**, 16487 (1996).

[52] A. B. Migdal, JETP **7**, 996 (1958).

[53] O. D. Restrepo, K. Varga, and S. T. Pantelides, App. Phys. Lett. **94**, 212103 (2009).

[54] W. L. McMillan, Phys. Rev. **167**, 331 (1968).

[55] P. B. Allen and R. C. Dynes, Phys. Rev. B **12**, 905 (1975).

[56] M. Lüders, M. A. L. Marques, N. N. Lathiotakis, A. Floris, G. Profeta, L. Fast, A. Continenza, S. Massidda, and E. K. U. Gross, Phys. Rev. B **72**, 024545 (2005).

[57] M. A. L. Marques, M. Lüders, N. N. Lathiotakis, G. Profeta, A. Floris, L. Fast, A. Continenza, E. K. U. Gross, and S. Massidda, Phys. Rev. B **72**, 024546 (2005).

[58] G. Profeta, C. Franchini, N. N. Lathiotakis, A. Floris, A. Sanna, M. A. L. Marques, M. Lüders, S. Massidda, E. K. U. Gross, and A. Continenza, Phys. Rev. Lett. **96**, 047003 (2006).

[59] J. M. Ziman, *Electrons and Phonons* (Oxford University Press, New York, 2001).

[60] P. B. Allen, Phys. Rev. B **13**, 1416 (1976).

[61] P. B. Allen, Phys. Rev. B **17**, 3725 (1978).

[62] B. Xu and M. J. Verstraete, arXiv:cond-mat p. 1311.6805 (2013).

[63] P. B. Allen and V. Heine, J. Phys. C: Sol. State Phys. **9**(12), 2305 (1976).

[64] P. B. Allen and M. Cardona, Phys. Rev. B **23**, 1495 (1981).

[65] F. Giustino, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **105**, 265501 (2010).

[66] S. Poncé, G. Antonius, P. Boulanger, E. Cannuccia, A. Marini, M. Côté, and X. Gonze, Comp. Mat. Sci. **83**, 341 (2014).

C 3 Molecular Dynamics Simulations [1]

Riccardo Mazzarello
Institute for Theoretical Solid State Physics
RWTH Aachen

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Molecular dynamics is a computer simulation of the motion of interacting particles (atoms or molecules). In this method, particles are treated as classical objects evolving according to Newton's equations of motion. To include quantum effects on the particle motion, more sophisticated methods have been developed, e.g. path integral molecular dynamics, which will not be discussed here.

Molecular dynamics simulations can be used to compute the equilibrium thermodynamic variables of a system, which can be expressed as statistical (ensemble) averages. Evaluation of these quantities is possible under the assumption that the ergodic hypothesis [1, 2] holds. Roughly speaking, this hypothesis states that the ensemble average of a thermodynamic property A is equal to its time average:

$$\langle A(\mathbf{R}_I, \mathbf{P}_I)\rangle = \lim_{t\to\infty} \frac{1}{t} \int_0^t A(\mathbf{R}_I(\tau), \mathbf{P}_I(\tau))d\tau, \tag{1}$$

where $\langle ...\rangle$ denotes the relevant ensemble average and the right-hand side denotes the integral of A over a typical trajectory.

Molecular dynamics can also be employed to investigate the structural and dynamical properties of a system of particles. As an example of a dynamical quantity, we consider here the macroscopic diffusion coefficient D. Assuming the system is isotropic, this quantity can be easily computed from

$$D = \frac{1}{2d} \lim_{t\to\infty} \frac{d}{dt}\langle r^2(t)\rangle, \tag{2}$$

where d denotes the dimensionality of the system. This formula relates a macroscopic quantity, D, to the microscopic mean-squared distance $\langle r^2(t)\rangle$ over which the atoms have moved in the time interval t.

Molecular dynamics methods can be divided into several classes, which differ in how the forces describing the interactions between particles are modeled.

In *ab initio* molecular dynamics (also called first-principles molecular dynamics) [3, 4], the forces acting on the particles are derived from quantum-mechanical calculations of electronic energies. The most widely employed *ab initio* methods are based on the Born-Oppenheimer approximation. and on the derivation of the forces from electronic ground-state energies, using Hellmann-Feynman theorem [4]. The ground state energies are most often evaluated by density functional theory [5, 6]; however, other methods, such as Hartree-Fock theory, can also be employed to compute electronic ground-state properties, in principle.

In classical molecular dynamics, on the other hand, interactions between particles are modeled by classical potential functions [7] obtained by fitting against experiments or higher-level quantum-mechanical calculations. This scheme is computationally cheap but often less accurate than the *ab initio* one. In particular, standard classical potentials cannot describe changes in bonding and chemical reactions, nor they can describe metal binding and strong polarization effects in a satisfactory way. However, sometimes they are more accurate than *ab initio* methods! Consider, for instance, a system of particles (such as rare-gas atoms or non-polar molecules) interacting via van der Waals forces. Standard density functionals based on the local density approximation or the generalized gradient approximation generally fail to describe accurately long-range van der Waals forces [5]. On the other hand, the use of classical Lennard-Jones potentials,

$$E_{LJ}(R) = 4\epsilon[(\sigma/R)^{12} - (\sigma/R)^6], \tag{3}$$

where the parameters ϵ and σ are extracted from experimental data, often yields satisfactory results.

There are two other important molecular dynamics schemes, which lie between classical and *ab initio* methods. The first one employs semi-empirical potentials [8, 9]: these potentials are based on a quantum mechanical description of the electrons but several approximations are made to compute the matrix elements between the Hamiltonian and the basis set. More specifically, a) the most computationally demanding integrals are neglected and, b) to compensate for this, the remaining integrals are made into parameters and their values estimated from *ab initio* or experimental data. Tight-binding methods belong to this class of potentials.

The second scheme is based on the so called QM/MM methods [10, 11]: within this approach, a small part of the system is treated quantum-mechanically (the one where reaction phenomena occur) and the rest of the system is treated classically. The biggest challenge faced by these methodologies is how to model the boundary region and the interaction between quantum and classical parts. The QM/MM scheme is particularly suited to the study of large systems, in which chemical reactions occur in a small subsystem (e.g. a small portion of a protein interacting with an ion or a molecule). This scheme has been further generalized in the ONIOM method [12] by dividing the system into three or more parts and applying different *ab initio*, semi-empirical or classical methods to the different parts.

In the following sections we provide a concise introduction to molecular dynamics methods and explain how it works and why it is useful. The chapter is divided as follows: after providing some basic notions of statistical mechanics (section 2), we discuss how molecular dynamics works in some detail and then focus on the simulation of different statistical ensembles (section 3). The second part of the chapter is devoted to the two most important classes of potentials used in molecular dynamics simulations: classical potentials (section 4) and first-principles potentials based on density functional theory calculations of electronic ground-state energies (section 5).

2 Basic notions of statistical mechanics

It is more natural to derive the basic laws of statistical mechanics within the framework of quantum mechanics. We follow this approach, but also provide the classical expressions for the most important quantities. Although statistical mechanics is a very well-established branch of physics, its foundations are still a subject of debate and a topic of research. Here we only discuss the standard approach based on the equal a priori probability postulate (we mostly follow Ref. [1] and [2]). This postulate asserts that, at equilibrium, an isolated system with fixed number of particles N, volume V and energy E is found with equal probability in each of its microstates with energy E. Here the term microstate refers to a microscopic configuration of the system, whereas a macrostate describes its macroscopic, thermodynamic properties. The so-called microcanonical ensemble is the probability distribution for the states of an isolated system. If we denote by $\Omega(N, V, E)$ the number of microstates with energy E, then the probability associated with each of these microstates is $1/\Omega$. The quantity $\Omega(N, V, E)$ is also called microcanonical partition function. Notice that, since we are considering systems with a very large but finite number of particles in a finite volume, the corresponding Hamiltonian has discrete eigenvalues. Nevertheless, the eigenvalue degeneracy will typically be huge. We define the entropy of the system as

$$S(N, V, E) \equiv k_{\mathrm{B}} \ln \Omega(N, V, E). \tag{4}$$

Classically, an isolated system is constrained on a constant energy surface; however, for computational convenience, it is better to consider an energy shell instead of a surface (see, for instance, Ref. [13]). The volume of the shell is:

$$\Omega_{\Delta E}(N,V,E) = \Sigma(N,V,E)\Delta E, \tag{5}$$

where ΔE is a small quantity and $\Sigma(N,V,E)$ is the area of the energy surface. The corresponding expression for the entropy

$$S(N,V,E) = k_{\rm B}\ln[\Omega_{\Delta E}(N,V,E)/h^{dN}]. \tag{6}$$

This "classical" formula contains the quantity h^{dN}, which is the volume of a quantum state in the $2dN$ dimensional phase-space. Hence, $\Omega_{\Delta E}(N,V,E)/h^{dN}$ yields the total number of states in the energy shell. In the case of indistinguishable particles, this quantity must be further divided by the factor $N!$, in that all the points in the phase space which differ only by a permutation of the particles correspond to the same quantum state. The presence of this factor solves Gibbs paradox. It can be easily shown that, in the thermodynamic limit,

$$N \to \infty, \;\; V \to \infty, \;\; N/V = \text{const.}, \tag{7}$$

it holds

$$\ln\Omega_{\Delta E}(N,V,E) \to \ln\Omega_{E}(N,V,E), \tag{8}$$

where $\Omega_E(N,V,E)$ is the total volume enclosed by the energy surface. Hence, in this limit the classical entropy can be written as

$$S(N,V,E) = k_{\rm B}\ln[\Omega_{E}(N,V,E)/h^{dN}]. \tag{9}$$

Now we assume that the isolated quantum system can be divided into two weakly-interacting systems, A and B, such that the two subsystems can exchange energy but have well defined energies E_A and E_B with $E = E_A + E_B$. We want to find the most likely distribution of energies. For a given energy E_A, the total number of degenerate states of the system is $\Omega_A(E_A)\times\Omega_B(E-E_A)$ or, in terms of the logarithm of $\Omega(E)$:

$$\ln\Omega(E_A, E-E_A) = \ln\Omega_A(E_A) + \ln\Omega_B(E-E_A). \tag{10}$$

To find the most likely value of E_A, we calculate the derivative of $\ln\Omega(E_A, E-E_A)$ with respect to E_A and set it to zero. We obtain from formula (10):

$$\left(\frac{\partial\ln\Omega_A(E_A)}{\partial E_A}\right)_{N_A,V_A} = \left(\frac{\partial\ln\Omega_B(E_B)}{\partial E_B}\right)_{N_B,V_B}. \tag{11}$$

When this condition is met, the two subsystems can be considered to be in equilibrium. Recalling the thermodynamic relation between temperature and entropy,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V}, \tag{12}$$

the meaning of formula (11) becomes clear: it simply states that, at equilibrium, the two subsystems have the same temperature. Here we just assume that the entropy and temperature defined

above are equivalent to the thermodynamic quantities we are familiar with, without proving it (for a more thorough discussion see, for instance, Refs. [1, 2]).
In the following we consider the case where the subsystem B is much larger than A, so that it can be considered as a heat bath. We assume that the volume and the number of particles of A remain constant. If the system A is in a microstate α with energy E_α, then the degeneracy of the heat bath is equal to $\Omega_B(E - E_\alpha)$. Hence, the probability to find the subsystem A in this microstate is:

$$P_\alpha = \frac{\Omega_B(E - E_\alpha)}{\sum_\beta \Omega_B(E - E_\beta)}, \tag{13}$$

where the sum is over all microstates of A. Since the subsystem B is much larger than A, we can safely assume that $E_\beta \ll E$ for each E_β. For the same reason, we can use the definition of entropy for an isolated system given in formula (4) and rewrite formula (13) as

$$P_\beta = \frac{\exp(S_B(E - E_\alpha)/k_\mathrm{B})}{\sum_\beta \exp(S_B(E - E_\beta)/k_\mathrm{B})}. \tag{14}$$

We now expand $S_B(E - E_\beta)$ around E:

$$S_B(E - E_\beta) \sim S_B(E) - E_\beta \frac{\partial S_B(E)}{\partial E} = S_B(E) - \frac{E_\beta}{T}. \tag{15}$$

Inserting this expansion into formula (14), we finally obtain

$$P_\alpha = \frac{\exp(-\frac{E_\alpha}{k_\mathrm{B}T})}{\sum_\beta \exp(-\frac{E_\beta}{k_\mathrm{B}T})}. \tag{16}$$

This is the probability distribution for a system at constant N, V and T, i.e. the canonical ensemble. The denominator is the partition function Z:

$$Z(N, V, T) \equiv \sum_\beta \exp\left(-\frac{E_\beta}{k_\mathrm{B}T}\right) = \sum_E \exp\left(-\frac{E}{k_\mathrm{B}T}\right) \Omega(N, V, E), \tag{17}$$

where the last equality provides a link between Z and the microcanonical partition function Ω. Assuming identical particles, the classical limit for the partition function reads:

$$Z(N, V, T) = \frac{1}{h^{dN} N!} \int_V d\mathbf{r}^N \int d\mathbf{p}^N \exp\left[-\frac{1}{k_\mathrm{B}T}\left(\sum_i \frac{p_i^2}{2M} + U\left(\{\mathbf{r}_i\}\right)\right)\right], \tag{18}$$

where p_i is the momentum of particle i, M is the mass of the particles and $U(\{\mathbf{r}_i\})$ is the potential energy. Again, the factor $1/N!$ accounts for the indistinguishability of the particles. The quantity h^{dN} makes Z into a dimensionless quantity. The volume dependence of the classical expression of Z is contained in the limits of integration for the coordinates $\mathbf{r}_i$. Notice that the integral over momenta can be easily carried out, yielding

$$Z = \frac{1}{N! \Lambda^{dN}} Z_c, \tag{19}$$

where $\Lambda = \sqrt{h^2/2\pi m k_\mathrm{B} T}$ is the thermal De Broglie wavelength and Z_c is the configurational partition function:

$$Z_c = \int_V d\mathbf{r}^N \exp\left(-\frac{1}{k_\mathrm{B}T} U\left(\{\mathbf{r}_i\}\right)\right). \tag{20}$$

We define the Helmholtz free energy as

$$F(N,V,T) \equiv -k_{\mathrm{B}}T \ln Z. \tag{21}$$

Since we now know the probability distribution for this ensemble, we can easily write down the expression for the average energy $\langle E \rangle$ of the system, which coincides with the thermodynamic internal energy U:

$$\langle E \rangle = \frac{\sum_\alpha E_\alpha \exp(-\frac{E_\alpha}{k_{\mathrm{B}}T})}{\sum_\beta \exp(-\frac{E_\beta}{k_{\mathrm{B}}T})} = \frac{\partial F/T}{\partial 1/T}. \tag{22}$$

The latter equality is a well-known thermodynamic relation between F and U.
We now drop the assumption that the volume of the subsystem A remains constant and assume that B is a constant pressure bath, besides being a heat bath. The number of particles N of A is fixed. We derive the probability distribution for the corresponding ensemble, which is called isothermal-isobaric ensemble. The derivation is analogous to the canonical case. The system $A+B$ is isolated and has constant energy and volume. The probability to find the subsystem A in the microstate α having energy E_α and volume V_α is:

$$P_\alpha = \frac{\Omega_B(E-E_\alpha, V-V_\alpha)}{\sum_\beta \Omega_B(E-E_\beta, V-V_\beta)}, \tag{23}$$

where the sum is over all microstates of A. We can rewrite formula (23) in terms of entropies as well,

$$P_\alpha = \frac{\exp(S_B(E-E_\alpha, V-V_\alpha)/k_{\mathrm{B}})}{\sum_\beta \exp(S_B(E-E_\beta, V-V_\beta)/k_{\mathrm{B}})}, \tag{24}$$

and expand around E and V, taking into account that $E_\beta \ll E$ and $V_\beta \ll V$.

$$S_B(E-E_\beta, V-V_\beta) \sim S_B(E,V) - E_\beta \frac{\partial S_B(E,V)}{\partial E} - V_\beta \frac{\partial S_B(E,V)}{\partial V} \tag{25}$$

$$= S_B(E,V) - \frac{E_\beta}{T} - \frac{PV_\beta}{T}, \tag{26}$$

where P is the pressure of the system (not to be confused with the probabilities P_β). We insert this expansion in formula (24) and obtain the relevant probability distribution:

$$P_\alpha = \frac{\exp(-\frac{E_\alpha}{k_{\mathrm{B}}T} - \frac{PV_\alpha}{k_{\mathrm{B}}T})}{\sum_\beta \exp(-\frac{E_\beta}{k_{\mathrm{B}}T} - \frac{PV_\beta}{k_{\mathrm{B}}T})}. \tag{27}$$

The denominator is the isothermal-isobaric partition function, $\Delta(N,T,P)$:

$$\Delta(N,T,P) \equiv \sum_\beta \exp\left(-\frac{E_\beta}{k_{\mathrm{B}}T} - \frac{PV_\beta}{k_{\mathrm{B}}T}\right). \tag{28}$$

Equivalently, we can rewrite Δ as an integral over volumes and, for each volume, as a sum over all the corresponding microstates $\beta(V)$ of A:

$$\Delta(N,T,P) = \frac{1}{V_0}\int_V dV \sum_{\beta(V)} \exp\left(-\frac{E_{\beta(V)}}{k_{\mathrm{B}}T} - \frac{PV}{k_{\mathrm{B}}T}\right), \tag{29}$$

where V_0 is a constant having the unit of volume. Δ can be easily expressed in terms of the canonical and microcanonical partition functions:

$$\Delta(N,T,P) = \frac{1}{V_0}\int_V dV \exp\left(-\frac{PV}{k_\mathrm{B}T}\right) Z(N,V,T) \tag{30}$$

$$= \frac{1}{V_0}\int_V dV \sum_E \exp\left(-\frac{E}{k_\mathrm{B}T} - \frac{PV}{k_\mathrm{B}T}\right) \Omega(N,V,E). \tag{31}$$

The classical expression for Δ (assuming identical particles) is:

$$\Delta(N,T,P) = \frac{1}{V_0 h^{dN} N!}\int_0^\infty dV \int_V d\mathbf{r}^N \int d\mathbf{p}^N \exp\left[-\frac{1}{k_\mathrm{B}T}\left(\sum_i \frac{p_i^2}{2M} + U(\{\mathbf{r}_i\}) + PV\right)\right]. \tag{32}$$

We define the Gibbs free energy G as

$$G(N,T,P) \equiv -k_\mathrm{B}T \ln \Delta(T,P). \tag{33}$$

To show that this quantity indeed coincides with the thermodynamic potential which determines the behaviour of a system at constant T and P, one calculates Δ within the maximum term approximation (which is an excellent approximation for large systems):

$$\Delta(N,T,P) = \frac{1}{V_0}\int_V dV \sum_E \exp\left(-\frac{E}{k_\mathrm{B}T} - \frac{PV}{k_\mathrm{B}T}\right) \Omega(N,V,E) \tag{34}$$

$$= \frac{1}{V_0}\int_V dV \sum_E \exp\left(-\frac{E}{k_\mathrm{B}T} - \frac{PV}{k_\mathrm{B}T} + \frac{S}{k_\mathrm{B}}\right) \tag{35}$$

$$\sim \exp\left[-\frac{\langle E\rangle}{k_\mathrm{B}T} - \frac{P\langle V\rangle}{k_\mathrm{B}T} + \frac{S(N,\langle V\rangle,\langle E\rangle)}{k_\mathrm{B}}\right]. \tag{36}$$

From this follows immediately that

$$G = \langle E\rangle - TS + P\langle V\rangle. \tag{37}$$

3 How molecular dynamics works

3.1 Integration of the equations of motion

In this section, we consider molecular dynamics simulations of isolated systems, where the total number of particles N, the volume V and the total energy E are kept constant. For obvious reasons, these simulations are called NVE simulations. In principle, molecular dynamics appears to be the ideal computational tool to investigate isolated systems. However, one must keep in mind that energy is never strictly conserved during a run, due to discretization and numerical errors. To perform reliable NVE simulations, it is thereby crucial to employ algorithms for the integration of Newton's equations of motion which conserve energy to a good accuracy on long time scales.

A proper discretization of Newton's equations is clearly a prerequisite for a reliable and accurate simulation. The timestep Δt controls this discretization: it is the time interval between two consecutive evaluations of the forces. In principle, one would like to choose the timestep as

large as possible, for the calculation of forces is, generally, the most time-consuming task. However, to avoid large discretization errors, Δt must be smaller than the inverse of the fastest vibrational frequency ω_{max} in the system:

$$\Delta t < 1/\omega_{max}. \tag{38}$$

For this reason, typical timesteps used in molecular dynamics simulations are of the order of fs (for Car-Parrinello molecular dynamics, they must be significantly smaller: see section 5.2).
Turning now to integration algorithms, one of the simplest and most widely employed scheme is the Verlet algorithm [14],

$$\mathbf{R}_I(t + \Delta t) = 2\mathbf{R}_I(t) - \mathbf{R}_I(t - \Delta t) + \frac{\Delta t^2}{M_I}\mathbf{F}_I(t). \tag{39}$$

This algorithm turns out to work remarkably well, in spite of its simplicity. In particular, it shows small energy drifts on large time scales, due to the fact that a) it is time-reversible and b) it is symplectic, i.e. it preserves area in the phase space. Obviously, this property does not imply that the Verlet algorithm predicts accurately trajectories for long times. In fact, there exists no algorithm capable of doing that! Generally, the trajectory of a N-particle system depends sensitively on the initial conditions: two trajectories that are very close at time $t = 0$ will diverge exponentially as a function of time. This behaviour is called Lyapunov instability [15]. In particular, the trajectories generated in a simulation will differ slightly from the true ones at the beginning of the run due to intrinsic integration errors of the algorithms employed: the exponential divergence with respect to the true trajectory is thus unavoidable. Nevertheless, molecular dynamics is still useful because

- its goal is to investigate the statistical properties of a system, not to calculate its exact evolution starting from precisely known initial conditions;
- there is numerical evidence that there exist true trajectories, called shadowing (or shadow) orbits, which are similar to the numerical trajectories for long times [1, 16]. The latter property can be seen as a consequence of the shadowing lemma of classical mechanics, which states that systems experiencing slightly different potential surfaces must exhibit a similar dynamics as a whole, in spite of the exponential sensitivity of the trajectories to the initial conditions and/or potential energy differences. Hence, such systems display similar trajectories (possibly corresponding to different initial conditions), which remain close to each other on time scales which are long compared to the typical times needed for the Lyapunov instabilities to develop [16]. Since numerical trajectories are expected to follow closely some real trajectories for long time scales, they can be used to carry out statistical investigations of the behaviour of the system.

There are a few other popular algorithms, which are essentially equivalent to the Verlet algorithm (i.e. they yield the same trajectories) but differ in how velocities are evaluated (in the case of the Verlet algorithm, velocities are not explicitly computed), namely the Leap Frog algorithm, the Beeman scheme and the velocity-corrected Verlet algorithm. The latter two methods provide a more accurate evaluation of the velocities.
There exists another class of algorithms, called predictor-corrector methods, which are widely used to solve differential equations numerically. As far as the integration of Newton's equations is concerned, these methods have both advantages and disadvantages with respect to the simpler

Verlet algorithm [17]. The most important advantage is the possibility to use longer time steps without sacrificing the short-time energy conservation. On the other hand, predictor-corrector methods are usually neither time-reversible nor symplectic: these drawbacks generally lead to energy dissipation on long time scales. In the last section, we will discuss a predictor-corrector scheme which, although slightly dissipative, enables one to perform extremely efficient *ab initio* molecular dynamics simulations.

3.2 Thermostats

Thermostats are needed to perform simulations at constant temperature, including investigation of canonical ensembles (NVT simulations) and isobaric-isothermal ensembles (NPT simulations). Since, at thermal equilibrium, the equipartition theorem holds, every atom has an average kinetic energy equal to $3/2k_{\mathrm{B}}T$. This property is exploited to calculate the temperature of the system as a function of time in molecular dynamics simulations:

$$T(t) = \frac{1}{3k_{\mathrm{B}}N} \sum_{i=1}^{N} M_i v_i^2(t), \tag{40}$$

where v_i is the velocity of particle i. One has to keep in mind that, for a finite canonical system, the temperature T is not strictly constant but fluctuates, because the kinetic energy per particle fluctuates. If we keep the kinetic energy per particle constant, we do not simulate the true canonical ensemble but another ensemble, called isokinetic ensemble [1].
Thermostats are employed to add or remove energy to the system, so as to keep the average temperature constant. An ideal thermostat should satisfy the following requirements:

- sample the canonical distribution;
- do not perturb the particle dynamics significantly;
- be easy to implement and use, in particular be tunable via a coupling parameter.

The first point may sound like an obvious requirement, nevertheless it turns out that many widely employed thermostats do not have this property. In spite of this shortcoming, such thermostats can yield reliable results for sufficiently large systems, in that, in the thermodynamic limit, the average properties of a system do not depend on the ensemble selected.
Thermostats can be divided into two categories: stochastic and deterministic thermostats. In the case of stochastic thermostats, random variables are used to mimic the interaction of the system with the heat bath. Examples of stochastic thermostats are Langevin thermostats [18], the Andersen [19] thermostat and the recently developed Bussi [20] thermostat. On the contrary, deterministic thermostats, such as Berendsen [21] thermostat and Nosé-Hoover [22, 23] thermostats, do not use any random quantity. In the following, we briefly discuss the properties and the behaviour of selected thermostats in more detail.
- Velocity rescaling
The velocities of all the particles are multiplied by the same factor α, calculated by enforcing the total kinetic energy K to be equal to the average one, $\langle K \rangle$, at the target temperature: $\alpha = \sqrt{\langle K \rangle / K}$. Rescaling can be carried out at a selected frequency or when K exceeds a threshold value. This thermostat does not sample the canonical ensemble.

- Andersen thermostat
It is a simple stochastic thermostat which correctly samples the canonical ensemble. At each step, a prescribed number of particles are selected randomly and their velocities are drawn from the appropriate Maxwell-Boltzmann distribution at the chosen temperature T:

$$P(\mathbf{v}) = \left(\frac{M}{2\pi k_{\mathrm{B}} T}\right)^{3/2} \exp\left(-\frac{M v^2}{2 k_{\mathrm{B}} T}\right). \tag{41}$$

This prescription is meant to mimic collisions with the particles of the bath at temperature T. This thermostat is not very efficient though and introduces discontinuities in the trajectories.
- Langevin thermostat
It is a stochastic thermostat which applies a friction force $\gamma \mathbf{v}_I$ and a random force $\mathbf{\Xi}$ to each particle, so that they evolve according to a Langevin dynamics:

$$M_I \frac{d^2}{dt^2}\mathbf{R}_I = \mathbf{F}_I - \gamma \frac{d}{dt}\mathbf{R}_I + \mathbf{\Xi}, \tag{42}$$

where $\mathbf{\Xi}$ obeys

$$\langle \Xi_i(0)\Xi_j(t)\rangle = 2\gamma k_{\mathrm{B}} T \delta_{i,j}\delta(t). \tag{43}$$

This thermostat samples the canonical ensemble but strongly affects the dynamics of the particles. Hence, it should not be employed to investigate the dynamical properties of a system.
- Berendsen thermostat
The equations of motion are supplemented by a first-order equation for the kinetic energy K, whose driving force is the difference between the instantaneous kinetic energy and its target value:

$$\frac{dK}{dt} = \frac{\langle K\rangle - K}{\tau}, \tag{44}$$

where τ is the relaxation time of the thermostat. τ is a tunable parameter, which controls the strength of the coupling between the system and the thermostat: the larger τ, the weaker the coupling. This thermostat is deterministic, stable and easy to use but does not sample the canonical ensemble. Since it can steadily drive the system towards the equilibrium state, it is often used to equilibrate the system at the start of the simulation.
- Nosé-Hoover thermostats
The Lagrangian of the system is extended by introducing an additional degree of freedom, which describes the heat bath. The equations of motion obeyed by this generalized coordinate guarantee that the true degrees of freedom of the system sample a canonical ensemble. These thermostats sometimes show non-ergodic behaviour: to compensate for this drawback, chains of Nosé-Hoover thermostats have been introduced.
- Bussi thermostat
This thermostat is a generalization of both the velocity rescaling scheme and Berendsen thermostat. Within this approach, the scaling factor α is given by $\sqrt{K'/K}$, where K' is obtained by evolving K using a stochastic dynamics, which leaves the canonical distribution of K invariant. The stochastic dynamics contains a deterministic term, which coincides with Berendsen's driving force. This thermostat samples the canonical ensemble if the algorithm used to integrate the equations of motion is ergodic (e.g. the Verlet algorithm).

3.3 Barostats

Barostats are employed to carry out constant pressure simulations, e.g. to perform NPT simulations. In the following we provide a description of the most commonly employed barostats.

- Berendsen barostat
The Berendsen barostat [21] is similar in spirit to the Berendsen thermostat. Pressure is made to obey the equation of motion:

$$\frac{dP}{dt} = \frac{\langle P \rangle - P}{\tau_P}, \tag{45}$$

where τ_P is the relaxation time of the barostat. As a result, the size of the cell varies: more specifically, the volume of the cell is scaled by a factor η, which evolves according to

$$\eta(t) = 1 - \frac{\Delta t}{\tau_P}\beta_T(\langle P \rangle - P(t)), \tag{46}$$

where β_T is the isothermal compressibility of the system. Here the scaling is assumed to be isotropic (i.e. the shape of the cell does not change) but it is possible to consider anisotropic scaling as well.
- Andersen barostat
The Andersen barostat [19] is an extended system method, which bears some analogies with the Nosé-Hoover thermostat. The system is coupled to the the volume V of the simulation box, which is treated as an external variable. The coupling is supposed to mimic the action of a piston on the system. The piston has an effective mass Q, which is a tunable parameter. The kinetic and potential energy of the piston read:

$$E_{piston} = \frac{1}{2}Q\left(\frac{dV}{dt}\right)^2 + \langle P \rangle V. \tag{47}$$

The potential energy stems from the external pressure (equal to the target pressure $\langle P \rangle$) acting on the piston. Within this method, it is not possible to change the shape of the cell during the simulation.
- Parrinello-Rahman barostat
The Parrinello-Rahman barostat [24] is an extension of Andersen barostat, which allows one to change both the shape and the size of the cell in a simulation. This method has proven to be particularly useful in the study of structural phase transitions in solids. The simulation cell can be described in full generality by the three vectors $\mathbf{a}, \mathbf{b}, \mathbf{c}$, which represent the sides of the cell. If we define the (3×3) matrix $H \equiv \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$, then the volume of the cell is given by $V = \det H = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$. The kinetic and potential energies of the fluctuating volume are

$$E_{volume} = \frac{1}{2}Q\mathrm{Tr}\left(\frac{dH^T}{dt}\frac{dH}{dt}\right) + \langle P \rangle V. \tag{48}$$

3.4 Computation of equilibrium properties

Some thermodynamic quantities can be expressed in terms of a statistical average. Such quantities can be calculated directly from a molecular dynamics simulation, using the ergodic hypothesis. For instance, the internal energy of a system at constant T and V is given by a canonical average (Formula (22)), which can be estimated during a simulation by calculating a time average

$$\langle E \rangle = \lim_{t \to \infty} \frac{1}{t} \int_0^t E d\tau. \tag{49}$$

Time averages must be taken over sufficiently long times, so as to allow the system to explore the relevant portion of the phase space. One should keep in mind, however, that the ergodic

hypothesis is not universally true: there are systems, such as nearly harmonic solids, which exhibit ergodicity breaking. Furthermore, there are classes of systems for which an astronomically large time would be needed to explore the phase space, e.g. systems trapped in a metastable phase, such as glasses, or systems in a symmetry-broken phase, such as ferromagnets.
There are thermodynamic variables which cannot be expressed in terms of a statistical average and, thereby, cannot be computed from a molecular dynamics run in a straightforward way. Examples of such quantities are the entropy S and the Helmholtz free energy F, which are expressed in terms of the microcanical and canonical partition function respectively, see Formulas (4) and (21). Hence, to compute F or S, a different strategy must be adopted. In the following we focus on F. It turns out that the derivative of F with respect to specific thermodynamics variables is given by a statistical average. For instance, the derivative of F/T with respect to $1/T$ (at constant V) is equal to the energy. This suggests a route to calculate F based on the integration of some derivative of F along a path which links the state we are interested in to another state, whose F is known. In simulations one can also consider the derivative of F with respect to some coupling parameter λ, on which the potential energy U is assumed to depend. This approach is called Kirkwood's coupling parameter method [25]. Here we follow this approach and assume that $U(\lambda = 0)$ and $U(\lambda = 1)$ correspond to the reference system and the system of interest, respectively. The derivative of F with respect to λ can be expressed as an ensemble average:

$$\left(\frac{\partial F}{\partial \lambda}\right)_{N,V,T} = -k_{\mathrm{B}}T\frac{1}{Z}\left(\frac{\partial Z}{\partial \lambda}\right) = \frac{\int d\mathbf{r}^N(\partial U/\partial\lambda)\exp[-U(\lambda)/k_{\mathrm{B}}T]}{\int d\mathbf{r}^N \exp[-U(\lambda)/k_{\mathrm{B}}T]} = \left\langle \frac{\partial U}{\partial \lambda}\right\rangle_{\lambda}, \qquad (50)$$

where $\langle \partial U/\partial\lambda\rangle_\lambda$ denotes the ensemble average of $\partial U/\partial\lambda$ for a system with potential energy $U(\lambda)$. The free energy can then be computed as:

$$F(\lambda = 1) = F(\lambda = 0) + \int_0^1 d\lambda \left\langle \frac{\partial U}{\partial \lambda}\right\rangle_{\lambda}. \qquad (51)$$

Usually, a linear function is used to connect the initial and final potential:

$$U(\lambda) = U(0) + \lambda(U(1) - U(0)). \qquad (52)$$

It is also of great interest to compute the free energy barriers separating different phases (typically, a stable and a metastable phase) of a system. In case of high barriers, however, standard molecular dynamics methods are of little use, in that the overcoming of such barriers brought about by thermal fluctuations is a rare event, which generally occurs on time scales much longer than those accessible by simulations.
Several enhanced sampling methods [26, 27, 28, 29, 30] have been developed, which, in combination with molecular dynamics or Monte Carlo simulations, enable one to compute the energy barriers separating the different phases. The thermodynamic integration method introduced above can also be used to calculate free energy barriers along a reaction path defined by a proper reaction coordinate s [31]. In this approach the thermodynamic integration variable is not an independent coupling parameter but the reaction coordinate itself, which is a function of the coordinates of the particles, $s = s(\{\mathbf{r}_i\})$. The free energy difference reads

$$F(s_2) - F(s_1) = \int_{s_1}^{s_2} ds' \left\langle \frac{\partial U}{\partial s'}\right\rangle_{s',\,\mathrm{cond}}, \qquad (53)$$

where the last subscript indicates that, in contrast to Formula (50), the integrand is a conditional average calculated at $s(\{\mathbf{r}_i\}) = s'$:

$$\left\langle \frac{\partial U}{\partial s} \right\rangle_{s',\,\mathrm{cond}} \equiv \frac{\int d\mathbf{r}^N (\partial U/\partial s) \exp[-U/k_\mathrm{B}T]\delta(s' - s(\{\mathbf{r}_i\}))}{\int d\mathbf{r}^N \exp[-U/k_\mathrm{B}T]\delta(s' - s(\{\mathbf{r}_i\}))}. \tag{54}$$

To calculate the conditional average efficiently, the so-called blue-moon ensemble method [32] can be applied: within this method, this quantity is evaluated by a time average over a constrained trajectory with s fixed at the selected value, $s(\{\mathbf{r}_i\}) = s'$.
Another important enhanced sampling technique is metadynamics [33, 34, 35]. This recently developed method is enjoying growing popularity due to its high degree of flexibility: metadynamics has been successfully applied to many different fields, including chemistry, biochemistry and biophysics, solid state physics and materials science. This method is also based on the assumption that the process under study can be described in terms of a small number of reaction coordinates called collective variables, which depend on the coordinates of the particles. The method amounts to adding a history-dependent bias potential V_b, typically consisting of a sum of repulsive Gaussian functions, in the coarse-grained space of the collective variables, thus accelerating the occurrence of rare events and enabling an efficient exploration of the free energy surface of the system.
Denoting with $\mathbf{s} = (s_1(\{\mathbf{r}_i\}), ..., s_d(\{\mathbf{r}_i\})$ the set of collective variables, the expression for the free energy in this coarse-grained space, $F(\mathbf{s})$, is obtained by integrating out the irrelevant degrees of freedom:

$$F(\mathbf{s}') = -k_\mathrm{B}T \ln\left(\int d\mathbf{r}^N \exp\left(-\beta U\left(\{\mathbf{r}_i\}\right)\right)\delta\left(\mathbf{s}' - \mathbf{s}\left(\{\mathbf{r}_i\}\right)\right)\right). \tag{55}$$

The probability distribution for the variables $\mathbf{s}$ can be expressed in terms of $F(\mathbf{s})$:

$$P(\mathbf{s}') = \frac{\exp[-F(\mathbf{s}')/(k_\mathrm{B}T)]}{\int \exp[-F(\mathbf{s})/(k_\mathrm{B}T)]d\mathbf{s}}. \tag{56}$$

If we consider, for simplicity, a 1D coarse-grained space, the bias potential V_b is given by

$$V_b(s(\{\mathbf{r}_i\}), t) = w \sum_{t'=\tau,2\tau,...<t} \exp\left(-\frac{\left(s(\{\mathbf{r}_i\}) - s_{t'}\right)^2}{2\sigma_s^2}\right), \tag{57}$$

where $s_{t'}$ is the value of the CV of the system taken at time t', w and σ_s are the height and width of the Gaussians respectively and τ is the deposition interval. If the size of the Gaussians is small compared to the size of the free energy wells corresponding to the relevant phases of the system, then the Gaussians will gradually fill up the basins and flatten the barriers, until the evolution of the collective variables becomes diffusive (see Fig. 1). In this limit, the biasing potential yields an estimate of the free energy surface:

$$\lim_{t\to\infty} -V_G(\mathbf{s}, t) \sim F(\mathbf{s}). \tag{58}$$

It can be shown that a relation similar to (58) holds rigorously for the average of V_G over independent metadynamics runs, $\langle V_G \rangle_{runs}$ [36]:

$$\lim_{t\to\infty} -\langle V_G(\mathbf{s}, t)\rangle_{runs} = F(\mathbf{s}). \tag{59}$$

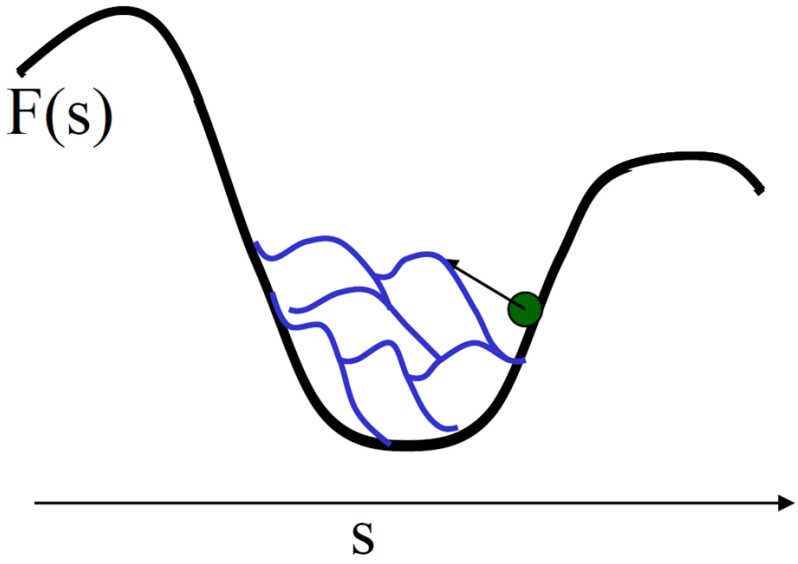


Fig. 1: *The history dependent potential fills up the free energy basin in the coarse-grained space of the collective variable s, thus allowing the system to escape the minimum.*

The accuracy and efficiency of metadynamics depend on the height w and the width σ_s of the Gaussians, as well as on the deposition interval τ [36]. To reduce the error, one can decrease the size of the Gaussians. However, the use of smaller Gaussians leads to an increase of the simulation times needed to fill the free energy basins. To improve the accuracy of the free energy reconstruction without sacrificing the efficiency of the method, new schemes have been proposed, wherein the Gaussian height [37] and/or width [38] are adapted to the local free energy or to the bias potential.

3.5 Computation of structural properties

Molecular dynamics allows one to investigate the equilibrium structural properties of a system. For disordered systems like liquids or glasses, the structural information provided by molecular dynamics is crucial to understand experimental results. Although liquids and glasses lack long-range crystalline order, the position of the particles are not truly random and some structural order is always present. This can be divided into short-range, medium-range and long-range order according to the typical length scales of the structural correlations [39].
In this respect, a very important quantity is the pair correlation function $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$, which describes the probability of finding any two particles in a small volume element about the points $\mathbf{r}_1$ and $\mathbf{r}_2$ respectively. In the canonical case, this quantity reads:

$$g^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{N(N-1)}{\rho^2} \frac{1}{Z_c} \int \exp(-\beta U(\{\mathbf{r}_i\})) d\mathbf{r}_3 \cdots \mathbf{r}_N, \tag{60}$$

where ρ is the density of the system.
For a homogeneous and isotropic system, $g^{(2)}$ does not depend on $\mathbf{r}_1$, nor on the direction of the vector $\mathbf{r}_1 - \mathbf{r}_2$. In this case, one can define a radial distribution function $g(r)$, which is a function of the distance $r = |\mathbf{r}_1 - \mathbf{r}_2|$ only. $g(r)$ describes the variation of the particle density as a function

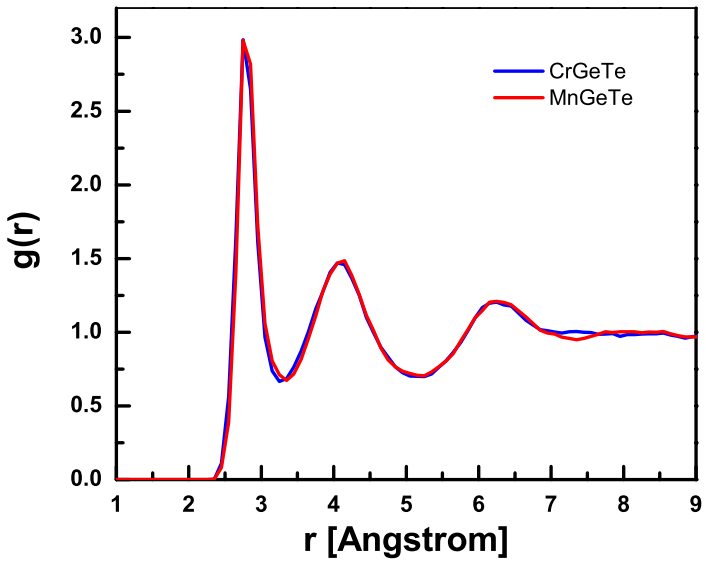


Fig. 2: *$g(r)$ of amorphous GeTe doped with Cr and Mn impurities calculated from an ab initio molecular dynamics run at T=300 K.*

of the distance r from a selected particle or, in other words, the probability of finding a particle at distance r from the reference particle (see Fig. 2). The function $g(r)$ provides information about the distribution of bond lengths. More precisely, the position of the first peak and second peak give the average interatomic distance between nearest neighbor and next nearest neighbour particles respectively. From these two quantities, one can calculate the average bond angle as well. It is more difficult to extract geometric information from higher order peaks. Another very important property of $g(r)$ is the fact that its Fourier transform, the so called structure factor $S(q)$, can be determined directly from neutron diffraction or x-ray diffraction experiments.

There are many other quantities one can compute from a molecular dynamics run, which yield information about the medium-range order (e.g. distribution of dihedral angles and primitive rings) and long-range order (e.g. Fourier transform of the atomic positions).

3.6 Computation of dynamical properties

As already mentioned in the introduction, molecular dynamics can also be used to calculate dynamical properties. Although these properties are related to non-equilibrium conditions, in the case of small perturbations they can be expressed in terms of equilibrium properties by using the relevant Green-Kubo relations [40].

In this section, we discuss how to calculate two important quantities characterizing the behaviour of liquids and glasses, namely the diffusivity D and the viscosity η. In the case of Stokes flow, these two quantities are related via the Stokes-Einstein relation,

$$D = \frac{k_B T}{6\pi\eta r}, \tag{61}$$

where r is the radius of the particle. One should keep in mind that this relation, which is based on macroscopic hydrodynamics, is not of universal validity though. The simplest way to calculate D is to use Formula (2) of the introduction. The mean-squared distance $\langle r^2(t)\rangle$ at time t is typically calculated by averaging over initial times and particles. An alternative way is to compute the velocity autocorrelation function, $\langle v_\alpha(t)v_\alpha(0)\rangle$ (here $v_\alpha(t)$ denotes the cartesian component $\alpha = x, y$ or z of the particle velocity at time t), and exploit the relation

$$D = \int_0^\infty \langle v_\alpha(t)v_\alpha(0)\rangle dt. \tag{62}$$

Formula (62) is an example of a Green-Kubo relation linking a transport coefficient to an autocorrelation function. The latter quantity is an equilibrium property of the system and can be calculated from a molecular dynamics run. More specifically, $\langle v_\alpha(t)v_\alpha(0)\rangle$ is also calculated by averaging over particles and initial times.

The viscosity η can also be evaluated using a Green-Kubo formula:

$$\eta = \frac{V}{k_\mathrm{B}T}\int_0^\infty \langle \tau_{\alpha\beta}(t)\tau_{\alpha\beta}(0)\rangle dt, \tag{63}$$

where $\tau_{\alpha\beta}$ denotes an off-diagonal element of the stress tensor. Generally, long simulation times (of the order of a few ns) are required to accumulate sufficient statistics to get converged values of η using Formula (63). This fact makes the calculation of η by *ab initio* molecular dynamics prohibitively expensive computationally. The statistical uncertainty of the data can slightly be reduced by averaging over independent off-diagonal tensor elements.

So far, we have not considered finite size effects. It turns out, however, that the use of periodic boundary conditions affects the diffusion coefficient considerably, if the size of the simulation cell is small. This problem can be severe in the case of *ab initio* simulations, where typical system sizes are of the order of hundreds of atoms. The dependence of D on the system size is due to the hydrodynamic interactions of a moving particle with the periodic images of the particle itself and of the other particles. Assuming that the simulation cell is a cube with size L and the liquid obeys Stokes equations, the following scaling relation holds [41, 42],

$$D(L) = D - \frac{k_\mathrm{B}T\xi}{6\pi\eta L}, \tag{64}$$

where $D(L)$ and D are the diffusivities of the finite and infinite system respectively and $\xi \sim 2.837297$. Since the right-hand side of this formula depends on η, it is possible to evaluate both D and η by doing a series of calculations of $D(L)$ for different L and fitting the data. Figure 3 shows the behaviour of $D(L)$ as a function of L for a simulation of CO_2 employing classical potentials.

As a last remark, we want to point out that, generally, thermostats affect the dynamical properties of a system significantly. This problem can be particularly severe for stochastic thermostats (with the exception of Bussi thermostat [20]). Hence, the most sensible way to compute dynamical properties at a given temperature is to perform microcanonical simulations, after equilibrating the system.

4 Classical molecular dynamics

Classical molecular dynamics is based on a description of the forces between particles in terms of empirical potential functions. In chemistry and biology, the empirical potentials are called

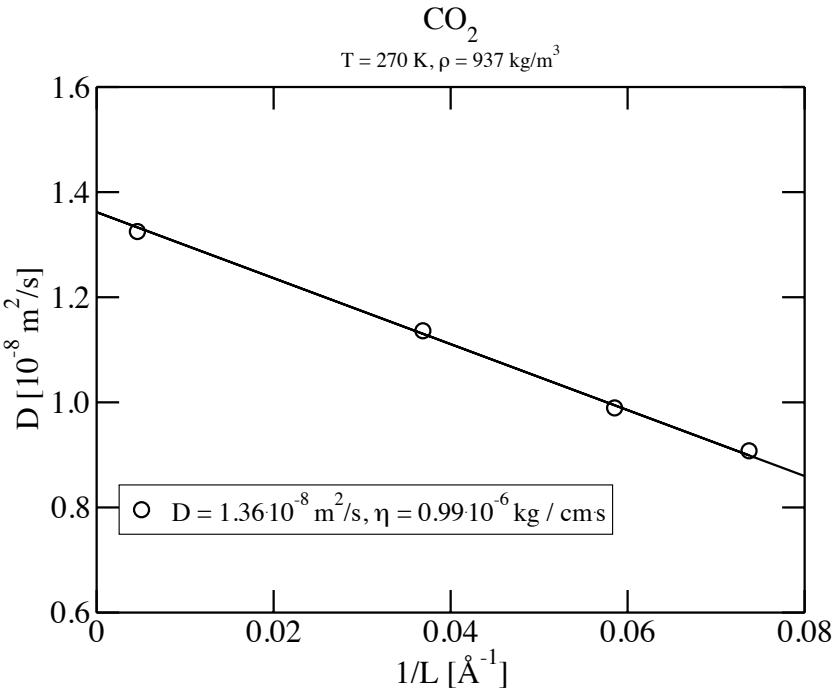


Fig. 3: *Transport properties of CO_2 obtained from classical molecular dynamics simulations. The interaction between molecules was modeled with a Lennard-Jones potential plus an electrostatic point-charge interaction (see next section).*

force fields. Force fields stand on the assumption that molecules are composed of units (such as bonds or, more generally, functional groups) which behave in a similar way in different chemical environments. Force fields consist of a) bonded terms describing the behaviour of the chemical bonds and b) non-bonded terms. The bonded terms typically consist of stretching potentials (two-body terms), bending potentials describing the bending of the angle formed by two bonds (three-body terms), torsional potentials describing the rotation around bonds (four-body terms) and cross terms which couple the previous three terms. The non-bonded terms consist of electrostatic interactions and van der Waals forces. Since these terms involve interactions between all the particles of the system, they are computationally more demanding than bonded terms. The electrostatic interactions are usually modeled by either assigning a point charge to each atom or by assigning a dipole moment to each bond. The relevant charges or moments are obtained by fitting against experimental or theoretical data. The van der Waals terms consist of a) a strong repulsive term at short distances due to the overlap of the electron clouds and b) an attractive term at large distances, due to induced dipole-dipole interactions, which varies as the sixth power of the distance. In the Lennard-Jones potential the repulsive term varies as R^{-12}: this is a computationally convenient choice.

Classical potentials consisting only of non-bonded terms can be used to simulate systems of weakly-interacting atoms (such as rare-gas atoms, for which only the van der Waals interaction must be considered) or rigid molecules (whose internal degrees of freedom are assumed to be frozen).

A severe shortcoming of standard force fields is the inability to describe the formation or breaking of chemical bonds. Recently, more sophisticated potentials, such as reactive force fields [43, 44] and neural network potentials [45], have been developed to overcome this problem.

Standard pair potentials or force fields do not work well for bulk metals: two-, three- and four-body potentials are not sufficiently accurate for these systems. To describe metallic cohesion, a large number of many-body potentials [46, 47], including, very recently, neural network potentials [48], have been developed.

5 *Ab initio* molecular dynamics

In the next two subsections, we introduce the two most popular *ab initio* molecular dynamics methods: Born-Oppenheimer (BO) molecular dynamics and Car-Parrinello (CP) molecular dynamics. Historically, the work by Car and Parrinello published in 1985 [3] marked the beginning of the *ab initio* molecular dynamics era. BO molecular dynamics became competitive with the CP method many years later [49]. Since BO molecular dynamics is conceptually simpler than the CP method, we discuss the former method first. Finally, in the last section, we introduce a novel molecular dynamics scheme, which is proving to be more efficient than both BO and CP methods.

5.1 Born-Oppenheimer molecular dynamics

BO molecular dynamics is based on the BO approximation, which roughly states that, since nuclei are much heavier than electrons, the latter can be assumed to follow the nuclei adiabatically, such that they are in their ground state at each nuclear configuration. Hence, in this approximation the electronic Hamiltonian reads:

$$[\hat{T} + \hat{V}_{ee}(\{\mathbf{r}_i\}) + \hat{V}_{ext}(\{\mathbf{r}_i\} : \{\mathbf{R}_\alpha\})]\Psi_0(\{\mathbf{r}_i\} : \{\mathbf{R}_\alpha\}) = E_0(\{\mathbf{R}_\alpha\})\Psi_0(\{\mathbf{r}_i\} : \{\mathbf{R}_\alpha\}), \quad (65)$$

where $\hat{T}$, $\hat{V}_{ee}$ and $\hat{V}_{ext}$ indicate the electronic kinetic energy, the electron-electron Coulomb interaction and the external potential operator due to the nuclei respectively. $\{\mathbf{r}_i\}$ and $\{\mathbf{R}_\alpha\}$ are the electronic and nuclear coordinates, E_0 and Ψ_0 are the electronic ground-state energy and wave function. E_0 and Ψ_0 depend parametrically on the positions of the nuclei. Ψ_0 must be computed at each time step in order to calculate the forces acting on the nuclei using Hellman-Feynman theorem [50]. If we define the total energy E_{tot} as

$$E_{tot} = E_0 + E_{nuclei}, \quad (66)$$

where E_{nuclei} is the repulsive interaction energy between nuclei, then the force acting on a nucleus α is given by

$$\mathbf{F}_\alpha = -\frac{\partial E_{tot}}{\partial \mathbf{R}_\alpha} = -\frac{\partial}{\partial \mathbf{R}_\alpha}\langle\Psi_0|\hat{H}|\Psi_0\rangle - \frac{\partial E_{nuclei}}{\partial \mathbf{R}_\alpha}. \quad (67)$$

Since Ψ_0 corresponds to the minimum of the energy functional $\langle\Psi|\hat{H}|\Psi\rangle$ (we assume that the wave functions are normalized), one gets

$$\mathbf{F}_\alpha = -\langle\Psi_0|\frac{\partial \hat{H}}{\partial \mathbf{R}_\alpha}|\Psi_0\rangle - \frac{\partial E_{nuclei}}{\partial \mathbf{R}_\alpha}. \quad (68)$$

If the external potential V_{ext} felt by the electrons is just the sum of the Coulomb potentials due to the nuclei, this expression can be further simplified:

$$\mathbf{F}_\alpha = -\int \rho_0(\mathbf{r})\frac{\partial V_{ext}(\mathbf{r}, \{\mathbf{R}\})}{\partial \mathbf{R}_\alpha} d\mathbf{r} - \frac{\partial E_{nuclei}}{\partial \mathbf{R}_\alpha}, \quad (69)$$

where ρ_0 is the electronic ground state density. If, on the other hand, the Hamiltonian contains effective non-local pseudopotentials [4] instead of the bare Coulomb terms, one has to use Eq. 68. Once forces are evaluated, integration of the classical equations of motion for the nuclei proceeds the same way as for classical molecular dynamics. In particular, the Verlet algorithm is used in most *ab initio* molecular dynamics codes.
To calculate E_0 is a formidable task, which cannot be solved exactly. As already pointed out in the introduction, the most popular method used to solve it approximately is, by far, density functional theory [5, 6]. This topic is covered in lecture A2 by M. Ležaić. Here we will just mention that, according to Hohenberg-Kohn theorem [51], E_0 is uniquely determined from ρ_0 and, for any given external potential V_{ext}, there exists a density functional $E_{V_{ext}}[\rho]$ such that

$$\min_{\rho} E_{V_{ext}}[\rho] = E_{V_{ext}}[\rho_0] = E_0. \tag{70}$$

Within the Kohn-Sham scheme [52], the ground-state density ρ_0 of the interacting system is evaluated by solving an auxiliary set of single-particle equations, called Kohn-Sham equations, self-consistently:

$$-\frac{\hbar^2}{2m}\nabla^2\varphi_i + (v_{ext} + v_H[\rho] + v_{xc}[\rho])\varphi_i = \epsilon_i\varphi_i, \tag{71}$$

where m is the mass of the electron, $v_H[\rho]$ is the Hartree potential and $v_{xc}[\rho]$ is the so called exchange-correlation potential. Both potentials depend on the total density of the system. The self-consistent solutions of Eq. 71 are called Kohn-Sham orbitals. They satisfy the property

$$\rho_0(\mathbf{r}) = \sum_{i=1}^{n} |\varphi_i(\mathbf{r})|^2, \tag{72}$$

where n denotes the number of occupied states.

5.2 Car-Parrinello molecular dynamics

We assume that Kohn-Sham orbitals φ_i are expanded in a normalized plane-wave basis set (the CP method is usually used in combination with plane waves, although other basis sets can also be employed, in principle):

$$\varphi_i(\mathbf{r}) = \sum_{\mathbf{G}} c_{i,\mathbf{G}} \exp(i\mathbf{G}\cdot\mathbf{r}), \tag{73}$$

For a given ionic configuration, the Kohn-Sham orbitals φ_i minimize the Kohn-Sham functional $E_{KS}[\{\varphi_i\}] \equiv E_{V_{ext}}[\rho[\{\varphi_i\}]]$ and the ground state energy is given by

$$E_0 = \min_{\varphi_i} E_{KS}[\{\varphi_i\}]. \tag{74}$$

If we express the orbitals in terms of the coefficients $c_{i,\mathbf{G}}$, then, for a finite set of plane waves, E_{KS} becomes a function of $c_{i,\mathbf{G}}$:

$$E_{KS}(\{c_{i,\mathbf{G}}\}) \equiv E_{KS}[\{\varphi_i(\{c_{i,\mathbf{G}}\})\}]. \tag{75}$$

The condition that the orbitals φ_i minimize the functional $E_{KS}[\{\varphi_i\}]$ and that the ground-state energy corresponds to the minimum of E_{KS} can be obviously restated as:

$$E_0 = \min_{c_{i,\mathbf{G}}} E_{KS}(\{c_{i,\mathbf{G}}\}). \tag{76}$$

Hence, the problem of finding E_0 can be seen as a classical minimization problem in the space of the coefficients $c_{i,\mathbf{G}}$. In other words, for a fixed nuclear configuration, the quantity $E_{KS}(\{c_{i,\mathbf{G}}\})$ can be considered as an effective potential energy defined in the configuration space spanned by the $c_{i,\mathbf{G}}$. Of course, this quantity contains both the kinetic energy and the potential energy of the quantum-mechanical electron gas.
There exists an effective method to tackle global optimization problems called simulated annealing [53]. A variant of this method based on molecular dynamics was used by Car and Parrinello [3] to minimize the Kohn-Sham functional. The method boils down to considering a fictitious dynamics for the coefficients $c_{i,\mathbf{G}}$ by introducing fictitious velocities $dc_{i,\mathbf{G}}/dt$:

$$\mu \frac{d^2}{dt^2} c_{i,\mathbf{G}} = -\frac{\partial E_{KS}}{\partial c^*_{i,\mathbf{G}}} - \sum_j \lambda_{ij} c_{j,\mathbf{G}}, \tag{77}$$

where μ is a fictitious electronic mass (with dimensions of mass times squared length) and λ_{ij} are Lagrange multipliers introduced to enforce the orthonormality of the wavefunctions. Suppose that the electronic system is trapped in a local minimum of $E_{KS}(c_{i,\mathbf{G}})$. Due to the fictitious kinetic energy, the system can overcome the surrounding barriers. Furthermore, it is possible to reach the global minimum (in principle) by gradually reducing the fictitious temperature down to zero.
The following step is to treat nuclear and fictitious electronic degrees of freedom on equal footing:

$$\mu \frac{d^2}{dt^2} c_{i,\mathbf{G}} = -\frac{\partial E}{\partial c^*_{i,\mathbf{G}}} - \sum_j \lambda_{ij} c_{j,\mathbf{G}}, \tag{78}$$

$$M_I \frac{d^2}{dt^2} \mathbf{R}_I = -\frac{\partial E}{\partial \mathbf{R}_I}, \tag{79}$$

where $E \equiv E_{KS}(\{c_{i,\mathbf{G}}\}, \{\mathbf{R}_I\}) + E_{nuclei}(\{\mathbf{R}_I\})$. By doing so, one can calculate the equilibrium nuclei configuration and achieve self-consistency simultaneously by simulated annealing.
One can also perform molecular dynamics simulations using the same classical equations of motion. The dynamics generated by these equations is generally different from the BO dynamics and therefore the forces on the ions are different from the BO ones: however, if the electronic degrees of freedom $c_{i,\mathbf{G}}$ remain close to the BO surface, i.e. if the fictitious electronic kinetic energy remains small, then the ionic trajectories will be good approximations to the BO ones. To keep the fictitious kinetic energy small, the following 3 conditions must be fulfilled.

1. Given the initial ionic configuration, the Kohn-Sham Hamiltonian must be solved self-consistently in the first molecular dynamics step to ensure that the electronic system is in its ground state at the initial time;

2. the initial fake velocities of the coefficients $c_{i,\mathbf{G}}$ must be small;

3. during the simulation, one has to make sure that the exchanges of energy between the nuclei and the fictitious electronic degrees of freedom are small. If this is the case, the fictitious kinetic energy will indeed remain small along the trajectory.

How can the third condition be achieved?

It turns out that the (fake) electronic and (true) ionic motion are decoupled if

- the system has an energy gap;
- the fictitious mass μ is sufficiently small.

In order to understand this point, it is useful to consider the limit of small deviations from the ground-state minimum, in which the electronic dynamics can be regarded as a superposition of harmonic oscillations: it can be shown (see Appendix A) that the corresponding electronic frequencies are given by [54]

$$\omega^1_{i,j} \sim (f_j(\epsilon^*_i - \epsilon_j)/\mu)^{1/2}, \tag{80}$$

$$\omega^2_{i,j} \sim ((f_i - f_j)(\epsilon_i - \epsilon_j)/2\mu)^{1/2}. \tag{81}$$

where ϵ^*_i and ϵ_j indicate the eigenvalue of the i-th unoccupied and j-th occupied level respectively. From these formulas follows that, if the system has a band gap, all the frequencies are finite. Furthermore, one can make the frequencies as large as needed by tuning μ. If all the fictitious electronic frequencies are significantly higher than the highest frequency of the ionic motion (which we denote as ω_I^{max}), then the two sets of oscillators, which approximately describe the motion of the two subsystems (ions and electrons), are decoupled. Hence, the basic physical fact behind the success of the method is that interacting oscillators with very different frequencies are practically decoupled!
How to choose the optimal value of the fictitious mass μ?

- On the one hand, one would like to set μ as small as possible, so as to make the minimum electronic frequency, $\omega_{i,j}^{min}$, much higher than ω_I^{max}.
- On the other hand, one would like the maximum electronic frequency, $\omega_{i,j}^{max}$, to be as small as possible. The reason for this is that the larger $\omega_{i,j}^{max}$ is, the smaller the time step Δt needs to be to describe the motion of the fictitious electronic degrees of freedom accurately (Δt must be smaller than $1/\omega_{i,j}^{max}$).

Since both $\omega_{i,j}^{min}$ and $\omega_{i,j}^{max}$ are approximately proportional to $\mu^{-1/2}$, a compromise must be made. Nevertheless, since it must hold $\omega_{i,j}^{max} > \omega_I^{max}$, the typical time steps used in CP molecular dynamics are smaller than those used in BO dynamics.
The ability of the CP method to keep the electron trajectories close to the BO surface is a consequence of classical adiabaticity, i.e. of the fact that the fake (classical) motion of the electronic coefficients $c_{i,\mathbf{G}}$ is much faster than the ionic motion. This is the classical counterpart of quantum adiabaticity: the latter yields the BO approximation. Two important results of the theory of adiabaticity in dynamical systems are relevant to CP molecular dynamics [54]:

1. a system of rapidly oscillating degrees of freedoms (the coefficients $c_{i,\mathbf{G}}$) perturbed by a slow variation of the parameters (the ionic coordinates R_I)) can be approximated by a system obtained by averaging over the fast motion;
2. there exist adiabatic invariants.

As far as point 1) is concerned, the motion of the coefficients $c_{i,\mathbf{G}}$ consists of rapid oscillations around the minimum of the Kohn-Sham energy functional and, by averaging over the fast motion, the trajectories of the ionic system closely resemble the BO ones. Notice that the fake

trajectories and the true trajectories (which lie on the BO surface) do not diverge over time. As a result, the CP forces, F_{CP}, will also oscillate around the values of the real BO forces, F_{BO}. Due to this property, there is no need to calculate forces very accurately, in contrast to the case of BO molecular dynamics.
The existence of adiabatic invariants is also a consequence of the fast dynamics of the electrons with respect to the ions. For a classical harmonic oscillator with a slowly varying frequency $\omega(t)$, the action $2\pi(E/\omega)$ where E is the total energy of the oscillator, is an adiabatic invariant, i.e. it is approximately conserved [55]. Hence, In CP dynamics, in the regime of small oscillations, the quantity $2\pi(E_\alpha/\omega_\alpha)$ is an adiabatic invariant for each normal mode α. Therefore, if the frequency $\omega_\alpha(t)$ is approximately equal to $\omega_\alpha(0)$, then $E_\alpha(t) \sim E_\alpha(0)$: there is (approximate) reversibility of the exchanges of energy between mode α and the rest of the system. Since, for linear oscillators, the average kinetic energy of a mode α is proportional to E_α, a metastable two-temperature regime can be maintained.
Typically, the Verlet algorithm is used to integrate both (fictitious) electronic and nuclear equations of motion:

$$c_{i,\mathbf{G}}(t+\Delta t) = 2c_{i,\mathbf{G}}(t) - c_{i,\mathbf{G}}(t-\Delta t) - \frac{\Delta t^2}{\mu}\left[\sum_{\mathbf{G}'} H_{\mathbf{G}\mathbf{G}'}c_{i,\mathbf{G}'}(t) - \sum_j \lambda_{ij}c_{j,\mathbf{G}}(t)\right], \quad (82)$$

$$\mathbf{R}_I(t+\Delta t) = 2\mathbf{R}_I(t) - \mathbf{R}_I(t-\Delta t) + \frac{\Delta t^2}{M_I}\mathbf{F}_I(t). \quad (83)$$

The first equation can be easily obtained from

$$\frac{\partial E}{\partial c^*_{i,\mathbf{G}}} = \sum_{\mathbf{G}'} H_{\mathbf{G}\mathbf{G}'}c_{i,\mathbf{G}'}. \quad (84)$$

We have mentioned that the CP method becomes problematic when there is no energy gap in the system, such as in metals. It is possible, however, to apply it successfully even to this class of systems either by performing periodic self-consistent energy minimizations [56] or by employing two thermostats [57]. In the latter case, the two thermostats are coupled to the ionic and the fake electronic degrees of freedom, respectively, and they are set at different T. Hence, the two subsystems can be kept at different T, so that the whole system remains in a metastable state. In other words, by using a separate thermostat for electrons, one can ensure that the electronic system does not heat up due to exchanges of energies with the ionic system, so that its fictitious kinetic energy remains very small.
As already mentioned, the CP method was the *ab initio* method of choice for several years, due to the two important advantages discussed above, namely the fact that there is no need to a) solve the Kohn-Sham Hamiltonian self-consistently at each time step and b) calculate forces very accurately. BO molecular dynamics started to become competitive with CP molecular dynamics in the 1990s [49], thanks to advances in algorithms to perform the minimization of the Kohn-Sham functional (e.g. improved conjugate-gradient and wave function extrapolation techniques) and the possibility to use larger time steps.

5.3 Recent developments

Recently, a very promising method has been developed by Kühne *et al*. [58], which combines the efficiency of the CP method with the large time steps used in BO molecular dynamics.

Within this method, no self-consistent diagonalization is performed during the simulation, however a predictor-corrector scheme is used to propagate the coefficients $c_{i,\mathbf{G}}$, instead of the CP equations of motion. More precisely, Kühne *et al*. employ the always stable predictor-corrector method introduced by Kolafa to study the classical dynamics of polarizable molecules [59]. This scheme makes it possible to maintain the system close to the BO surface, without having to use small time steps. So far, this method has been only implemented in the Quickstep program [60], where it is used in combination with a Gaussian, non-orthogonal basis set. Defining a $M \times N$ matrix $\mathbf{C}$ containing the expansion coefficients of the N occupied Kohn-Sham orbitals with respect to the M basis functions, the predictor-corrector scheme reads

$$\mathbf{C}^p(t_n) = \sum_{m=1}^{k}(-1)^{m+1}m\frac{\begin{pmatrix} 2K \\ K-m \end{pmatrix}}{\begin{pmatrix} 2K-2 \\ K-1 \end{pmatrix}}\mathbf{C}(t_{n-m})\mathbf{C}^T(t_{n-m})\mathbf{S}(t_{n-m})\mathbf{C}(t_{n-1}), \quad (85)$$

$$\mathbf{C}^c(t_n) = \omega \min[\mathbf{C}^p(t_n)] + (1-\omega)\mathbf{C}^p(t_n), \quad (86)$$

where $\omega = K/(2k-1)$ ($K \geq 2$), $\mathbf{S}$ is the $M \times M$ overlap matrix of the basis functions and $\min[\mathbf{C}^p(t_n)]$ denotes a single minimization step of the Kohn-Sham orbitals. This method is time reversible up to order h^{K+2} but it is not symplectic. Due to the latter property, the dynamics is dissipative and the method cannot generally be used to perform microcanonical simulations (although it is possible, in many cases, to make the dissipation very small by tuning the parameters of the method [58, 61]). To carry out canonical simulations, it is necessary to perform preliminary runs to estimate the intrinsic friction coefficient, γ^D, and then to add a random noise term, Ξ_I^D, so that the system obeys a Langevin-type equation (Eqs. 42 and 43).
It is also possible to add an additional friction term, $\gamma^L\mathbf{v}_I$, and the corresponding random noise term, Ξ_I^L. If $\gamma^L >> \gamma^D$, then the additional term dominates over the intrinsic term and there is no need to evaluate the latter term with great accuracy to sample the canonical ensemble.
Kühne's method has proven to be very efficient and to provide reliable results, irrespective of the band gap of the system [58]. Recent simulations of GeTe, $Ge_2Sb_2Te_5$ and similar small gap semiconductors have shown that the use of this method yields a factor of 20-30 speedup over conventional BO molecular dynamics [62, 63].

Appendices

A Fictitious dynamics of the electronic coefficients in the Car-Parrinello method

We follow Ref. [54]. Consider the expansion of a generic Kohn-Sham state φ_i with respect to an orthonormal basis:

$$\varphi_i = \sum_m c_{l,m}\zeta_m. \quad (87)$$

If we choose the basis set $\{\zeta_m\}$ to consist of the Kohn-Sham states φ_i themselves, then obviously the ground state of a system with n occupied states has coefficients

$$c^0_{l,m} = \begin{cases} \delta_{i,m} & m \leq n, \\ 0 & m > n. \end{cases} \quad (88)$$

If we now consider a state ψ_i "close" to φ_i, such that $\psi_i = \varphi_i + \delta\varphi_i$, then its coefficients can be expressed as

$$c_{i,m} = c^0_{i,m} + \delta c_{i,m}. \tag{89}$$

At first order in δc we have, for each i and j, due to the orthonormalization of the states,

$$\delta c_{i,j} + \delta c^*_{i,j} = 0. \tag{90}$$

Now we consider the matrix of Lagrange multipliers λ_{ij}. Since this matrix is Hermitian, it holds:

$$\lambda_{ij} = \frac{1}{2}(\lambda_{ij} + \lambda^*_{ij}). \tag{91}$$

From

$$\int d\mathbf{r}\ \psi^*_i(\mathbf{r},t)\psi_j(\mathbf{r},t) = \delta_{ij} \tag{92}$$

and

$$\mu\frac{d^2\psi_i(\mathbf{r},t)}{dt^2} = -f_i H\psi_i(\mathbf{r},t) + \sum_{j=1}^{n}\lambda_{ij}\psi_j(\mathbf{r},t), \tag{93}$$

follows

$$\begin{aligned}
\lambda_{ij} &= \frac{1}{2}(f_i+f_j)\int \psi^*_j H\psi_i + \frac{\mu}{2}\int \frac{d^2\psi_i}{dt^2}\psi^*_j + \frac{\mu}{2}\int \frac{d^2\psi^*_j}{dt^2}\psi_i && (94)\\
&= \frac{1}{2}(f_i+f_j)\int \psi^*_j H\psi_i - \mu\int \frac{d\psi_i}{dt}\frac{d\psi^*_j}{dt} && (95)\\
&= \frac{1}{2}(f_i+f_j)H_{ji} - \mu\int \frac{d\psi_i}{dt}\frac{d\psi^*_j}{dt}. && (96)
\end{aligned}$$

In the case of states ψ_i close to φ_i, the matrix elements $H_{ji} \equiv \int \psi^*_j H\psi_i$ of H between two occupied orbitals can be expressed as

$$H_{ij} = \sum_m \left(c^{0*}_{i,m} + \delta c^*_{i,m}\right)\left(c^0_{j,m} + \delta c_{j,m}\right)\epsilon_m. \tag{97}$$

The forces F_i acting on the perturbed orbitals ψ_i are given by

$$\begin{aligned}
F_i &= -\frac{\delta E}{\delta\psi_i} + \sum_{j=1}^{n}\lambda_{ij}\psi_j && (98)\\
&= -f_i H\psi_i + \sum_{j=1}^{n}\lambda_{ij}\psi_j && (99)\\
&\sim -f_i H\psi_i + \sum_{j=1}^{n}\left(\frac{f_i+f_j}{2}\right)H_{ji}\psi_j. && (100)
\end{aligned}$$

We have neglected the term $\mu \int (d\psi_i/dt)(d\psi_j^*/dt)$ of Eq. (94) because it is of second order in $\psi_i - \varphi_i$. We can now write explicitly the forces in terms of the coefficients $c_{i,j}^0$ and $\delta c_{i,j}$:

$$
\begin{aligned}
F_i &= -f_i \sum_m (c_{i,m}^0 + \delta c_{i,m})\epsilon_m \varphi_m + && (101)\\
&\quad \sum_{j=1}^{n} \left(\frac{f_i + f_j}{2} \right) \sum_m (c_{j,m}^{0*} + \delta c_{j,m}^*)(c_{i,m}^0 + \delta c_{i,m})\epsilon_m \sum_{m'} (c_{j,m'}^0 + \delta c_{j,m'})\varphi_{m'} && (102)\\
&= -\sum_{m=n+1}^{\infty} f_i(\epsilon_m - \epsilon_i)\delta c_{i,m}\varphi_m + && (103)\\
&\quad \sum_{j=1}^{n} \left(-f_i\epsilon_j \delta c_{i,j} + f_i\epsilon_i \delta c_{i,j} + \epsilon_j \delta c_{i,j} + \frac{f_i + f_j}{2}\epsilon_j \delta c_{i,j} + \frac{f_i + f_j}{2}\epsilon_i \delta c_{j,i}^* \right) \varphi_j. && (104)
\end{aligned}
$$

The last expression was obtained by discarding terms of second order in $\delta c_{i,j}$ and taking into account formula (88). We can use Eq. (90) to further simplify the expression for the forces. We finally obtain

$$
F_i = -\sum_{m=n+1}^{\infty} f_i(\epsilon_m - \epsilon_i)\delta c_{i,m}\varphi_m - \sum_{j=1}^{n} \frac{f_j - f_i}{2}(\epsilon_i - \epsilon_j)\delta c_{i,j}\varphi_j. \tag{105}
$$

From this formula, it follows immediately that

$$
\omega_{i,j}^1 \sim (f_j(\epsilon_i^* - \epsilon_j)/\mu)^{1/2}, \tag{106}
$$

$$
\omega_{i,j}^2 \sim ((f_i - f_j)(\epsilon_i - \epsilon_j)/2\mu)^{1/2}. \tag{107}
$$

References

[1] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, San Diego, 2002).

[2] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Butterworth-Heinemann, Oxford, 1996).

[3] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[4] R. Martin, *Electronic Structure: Basic Theory and Practical Applications* (Cambridge University Press, Cambridge, 2004).

[5] R. M. Dreizler and E. K. U. Gross, *Density Functional Theory* (Springer Verlag, Berlin, 1990).

[6] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, Oxford, 1989).

[7] F. Jensen, *Introduction to Computational Chemistry* (Wiley, Chichester, 2nd edition, 2007).

[8] J. Sadley, *Semi-Empirical Methods of Quantum Chemistry* (Wiley, New York, 1985).

[9] M. C. Zerner, Rev. Comput. Chem. **2**, 313 (1991).

[10] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).

[11] G. Monard and K. M. Merz, Jr., Acc. Chem. Res. **32**, 904 (1999).

[12] M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma, J. Phys. Chem. **100**, 19357 (1996).

[13] L. E. Reichl, *A Modern Course in Statistical Physics* (Wiley, New York, 2nd edition, 2004).

[14] L. Verlet, Phys. Rev. **159**, 98 (1967).

[15] A. M. Lyapunov, *The General Problem of the Stability of Motion* (Taylor & Francis, London, 1992).

[16] R. E. Gillilan and K. R. Wilson, J. Chem. Phys. **97**, 1757 (1992).

[17] G. Ciccotti and W. G. Hoover, *Molecular Dynamics Simulations of Statistical Mechanics Systems* (North-Holland, Amsterdam, 1986).

[18] T. Schneider and E. Stoll, Phys. Rev. B **17**, 1302 (1978).

[19] H. C. Andersen, J. Chem. Phys. **72**, 2384 (1980).

[20] G. Bussi, D. Donadio and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[21] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, J. Chem. Phys. **81**, 3684 (1984).

[22] S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[23] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[24] M. Parrinello and A. Rahman, Phys. Rev. Lett. **45**, 1196 (1980).

[25] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).

[26] G. M. Torrie and J. P. Valleau, J. Comput Phys. **23**, 187 (1977).

[27] P. A. Bash, U. C. Singh, F. K. Brown, R. Langridge and P. A. Kollman, Science **235**, 574 (1987).

[28] G. N. Patey and J. P. Valleau, J. Chem. Phys. **63**, 2334 (1975).

[29] F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).

[30] E. Darve and A. Pohorille, J. Chem. Phys. **115**, 9169 (2001).

[31] M. Sprik and G. Ciccotti, J. Chem. Phys. **109**, 7737 (1998).

[32] E. A. Carter, G. Ciccotti, J. T. Hynes and R. Kapral, Chem. Phys. Lett. **156**, 472 (1989).

[33] A. Laio and M. Parrinello, Proc. Natl Acad. Sci. USA **99**, 12562 (2002).

[34] A. Laio and F. L. Gervasio, Rep. Prog. Phys. **71**, 126601 (2008).

[35] A. Barducci, M. Bonomi and M. Parrinello, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **1**, 826 (2011).

[36] G. Bussi, A. Laio and M. Parrinello, Phys. Rev. Lett. **96**, 090601 (2006).

[37] A. Barducci, G. Bussi and M. Parrinello, Phys. Rev. Lett. **100**, 020603 (2008).

[38] D. Branduardi, G. Bussi and M. Parrinello, J. Chem. Theory Comput. **8**, 2247 (2012).

[39] S. Elliott, Nature **354**, 445 (1991).

[40] J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, London, 3rd edition, 2006).

[41] B. Dünweg and K. J. Kremer, J. Chem. Phys. **99**, 6983 (1993).

[42] I.-C. Yeh and G. Hummer, J. Phys. Chem. B **108**, 15873 (2004).

[43] A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard III, J. Phys. Chem. A **105**, 9396 (2001).

[44] T. T. Järvi, A. C. T. van Duin, K. Nordlund and W. A. Goddard III, J. Phys. Chem. A **115**, 10315 (2011).

[45] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007).

[46] M. S. Daw and M. I. Baskes, Phys. Rev. B **29**, 6443 (1984).

[47] F. Ercolessi, E. Tosatti and M. Parrinello, Phys. Rev. Lett. **57**, 719 (1986).

[48] N. Artrith, B. Hiller and J. Behler, Phys. Status Solidi B **250**, 1191 (2013).

[49] G. Kresse and J. Hafner, Phys. Rev. B **47**, 558 (1993).

[50] R. P. Feynman, Phys. Rev. **56**, 340 (1939).

[51] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[52] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[53] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, Science **220**, 671 (1983).

[54] G. Pastore, E. Smargiassi and F. Buda, Phys. Rev. A **44**, 6334 (1991).

[55] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, (Springer Verlag, Berlin, 2nd edition, 1989).

[56] I. Stich, R. Car and M. Parrinello, Phys. Rev. B **44**, 4262 (1991).

[57] P. E. Blöchl and M. Parrinello, Phys. Rev. B **45**, 9413 (1992).

[58] T. D. Kühne, M. Krack, F. R. Mohamed and M. Parrinello, Phys. Rev. Lett. **98**, 066401 (2007).

[59] J. Kolafa, J. Comput. Chem. **25**, 335 (2004).

[60] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, Comput. Phys. Commun. **167**, 103 (2005); www.cp2k.org.

[61] T. D. Kühne, M. Krack and M. Parrinello, J. Chem. Theory Comp. **5**, 235 (2009).

[62] S. Caravati, M. Bernasconi, T. D. Kühne, M. Krack and M. Parrinello, Appl. Phys. Lett. **91**, 171906 (2007).

[63] R. Mazzarello, S. Caravati, S. Angioletti-Uberti, M. Bernasconi and M. Parrinello, Phys. Rev. Lett. **104**, 085503 (2010).

C 4 Complex Magnetism[1]

Stefan Blügel
Peter Grünberg Institut
and
Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

1 Introduction to Complex Magnetism

Probably the most known manifestation of magnetism is ferromagnetism. Ferromagnetic materials are important constituents of many modern hi-tech devices. For example, more than 90% of all information world-wide is stored in ferromagnetic domains. Ferromagnetism is a form of magnetism, characterized at equilibrium by an additional macro-variable, the *spontaneous magnetization* $\boldsymbol{M}$, – a temporally stable *spontaneous collective ordering* of permanent magnetic moments typically of magnetic atoms – of a solid at temperatures below a critical temperature $T_{\rm c}$, called Curie temperature, $T_{\rm C}$, $T_{\rm C} = T_{\rm c}$, in case of ferromagnets. Above the critical temperature the spontaneous magnetization remains zero and the solid behaves like normal paramagnets. The disappearance of the spontaneous magnetization is an example for a phase transition of second order where the spontaneous magnetization is the order parameter characterizing the phase transition, vanishing continuously at $T_{\rm c}$. For second order phase transitions an important principal exists, namely *universality*. Universality means that the behavior of a system close to its critical point does not depend on details of the system such as its material parameters or the geometry of the sample. Instead, physical systems fall into so-called *universality classes* depending only on symmetry of the underlying model such as the spatial dimension of the system, or – in a magnetic system – the dimension of the spins. On the other hand the Curie temperature is a material dependent quantity, and a very important one. Since for technological applications magnetic devices must have operating temperatures of about room temperature or even higher, the understanding of the influence of the critical temperature on the detailed properties of the magnetic systems is an important part of today's research. This is discussed in much more detail in the lecture C1 of Phivos Mavropoulos.

For many magnets the permanent magnetic moments still exist above $T_{\rm c}$, but the magnetic moments are disordered. It may seem at first paradoxically that a disordered high-temperature phase is more symmetric than the low-temperature ordered state, but this is a quite general phenomenon. For example, a uniform liquid state (e.g. water) is invariant under arbitrary translations in space, while a crystal (e.g. ice) is only invariant under translations by an integer number of lattice constants. In fact, in most phase transitions some symmetry of the high-temperature phase becomes broken when an ordered state sets in. For example, ferromagnetic ordering breaks time reversal symmetry, i.e. if $T : t \longrightarrow -t$ then $\mathrm{M} \longrightarrow -\mathrm{M}$, because the magnetization M (average magnetic moment per unit volume) appearing below the transition temperature changes sign under time reversal. Such a symmetry breaking is called spontaneous, which means the Hamiltonian describing the system is invariant under a symmetry transformation, while the ordered state is not. See for example the Heisenberg exchange interaction, discussed in Section 2.1, which is invariant under time reversal symmetry since both $\mathbf{S}_i$ and $\mathbf{S}_j$ change sign when $t \longrightarrow -t$, while an ordered spin state with a nonzero average spin $\langle \mathbf{S} \rangle$ is not.

Magnetism is the science of cooperative effects of orbital and spin moments in matter, and ferromagnetism is one example of it. Actually all condensed matter shows a magnetic response to the presence of an applied magnetic field. The magnetizibility or susceptibility can vary greatly in both strength, by over 10 orders of magnitude, and manner. Some solids exhibit even a magneto-electric response, i.e. a magnetic response when an electric field is a applied. The response is used to classify solids for example in diamagnets, paramagnets or ferromagnets. Magnetism is a *quantum many-body phenomenon.* All-in-all, already in 1926, independent of each other, Dirac and Heisenberg recognized the quantum mechanical exchange interaction in competition to the kinetic energy of the moving electrons as the decisive origin of collective magnetism. The exchange interaction is a result of the electrostatic Coulomb interaction

between the electrons in concert with the antisymmetric nature of the many–electron wavefunction (Pauli principle). The strength of the exchange interaction is much modified by the electron correlation of the interacting electrons. More precisely, magnetism is a relativistic quantum phenomenon and their is probably no field of condensed matter where the spin-orbit interaction has so many different ramifications. Well-known and practically important manifestations of spin-orbit effects in magnets and spin-transport are the magnetic anisotropy, the anisotropic exchange, the antisymmetric exchange often referred to as Dzyaloshinskii-Moriya interaction, the tunnelling anisotropic magnetoresistance (TAMR), the anomalous and the spin-Hall effect, as well as their quantum counterparts, of the Elliot-Yafet, D'yakonov-Perel and Bir-Aronov-Pikus spin-relaxation mechanisms in electron transport, etc. For more details see also the lectures A6, A10, C6 of Yuriy Mokrousov, Gustav Bihlmayer, and Samir Lounis, respectively.

The quantum states of electrons vary tremendously from solid to solid if we consider that we deal with insulators, semiconductors, metals, oxides, molecular materials, crystalline, amorphous and glassy solids, liquids, heterostructures, thin films, quasi-one dimensional chains, clusters, single molecules giving rise to a rich spectrum of *exchange interactions* (for simplicity one refers frequently simply to: interactions between magnetic moments) connecting over distances the little permanent magnetic moments constituting the solid. We may deal with long- and short-ranged interactions, interactions that produce ferromagnetic or antiferromagnetic alignment of spins between different pairs of atoms, whose strength and alignment type may depend on the distance. This leads to a large variety of magnetic behaviors such as spin glass, spin liquids or spin ice and very different magnetic phases, e.g. *collinear magnetic phases* as for the ferromagnetic (FM), antiferromagnetic (AFM), ferrimagnetic (FIM), and sinusoidal spin-density wave phase (sSDW), or *non-collinear chiral and achiral magnetic phases* as for planar and conical helicoidal (HSDW) and cycloidal spin-density wave phase (CSDW) characterised by a single or multiple wave vector $\mathbf{q}$ states. (Please notice sSDW or SDW denotes the sinusoidal spin-density wave, while SSDW denotes the spiral spin-density wave, such as the HSDW and the CSDW.)

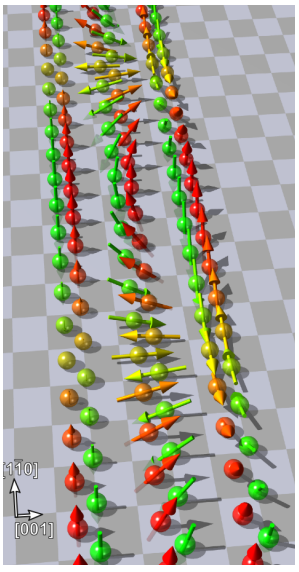


Fig. 1: *From left to right illustration of a single-q sinusoidal, helical and cycloidal spin-density wave of an AFM double chain (from [1]).*

Usually the full fermionic degree of freedom of the electronic quantum many-body problem is too large to discuss all these experimental observations effectively. Therefore, effective spin-models (see Section 2) have been developed that catch or try to catch the major microscopic exchange mechanism and reduce the problem to the spin-degree of freedom, by which the electron degree of freedom and the electronic structure is summarised by exchange parameters. One of the most familiar spin-models is the Heisenberg model (see Section 2.1 with the exchange interaction $J_{i,j}$ between atoms at sites i, and j. The presence of the spin-Hamiltonian bears various advantages, it permits (i) an efficient search for the magnetic ground-state, (ii) the calculations of dynamical properties such as the magnon dispersion or magnetic excitations and (iii) the calculation of thermal properties such as the critical temperature. In the last years there is an effort to calculate the exchange parameter from first-principles and enable the development of *realistic model Hamiltonians*, thus being able to model realistic systems.

In the past decade we witnessed a multitude of investigations of non-collinear magnetism in the context of first-principles theory, where solutions of both the spiral spin-density wave and

complex non-collinear magnetic structures have been found, first for simple, but now also for more complicated systems. Interesting steps are undertaken to explore the potential ground state magnetic structure, investigate the spin-dynamics or describe the thermodynamics properties in magnets. After the thorough discussion of density functional theory for a non-collinear magnet was completed in Ref. [2], the successful calculations established the density functional theory as a powerful tool to investigate these complex magnets, in particular complex itinerant magnets. For a deeper and broader introduction to the field following book and general review are recommended: [3, 4, 5, 6, 7].
In this Lecture, at first a brief introduction to the field of complex magnetism is given , and then an extension of the spin-density functional theory to non-collinear magnetism is introduced. Then, different strategies are provided in how to use the description of the non-collinear magnetism to find ground-states in complex magnets and the use of the adiabatic principle to calculated magnon and thermal properties. A slight emphasis is given to chiral magnetism, a field that enjoys currently considerable attention. Finally, some examples are discussed, taken from the field of magnetism in reduced dimension, i.e. thin films and clusters at surfaces, at which a lot of interesting new physics was observed recently.

1.1 Collinear magnetism

In collinear magnets there is a common magnetization axis $\hat{\mathbf{e}}_M$ for all atoms, for which for convenience the z-axis is chosen. In ferromagnets ($\uparrow\uparrow\uparrow\cdots$) the magnetic moment $M_i = M$ has the same value at all sites i, with the position of the atomic nucleus denoted by $\mathbf{R}_i$. An (Nèel) antiferromagnet ($\uparrow\downarrow\uparrow\cdots$) has a staggered arrangement of magnetic moments $M_i = \pm M$ of the same size, but opposite sign at neighboring sites. In real solids one can find, however, also more complicated forms of antiferromagnetism, e.g. the so-called "up-up-down-down" state ($\uparrow\uparrow\downarrow\downarrow\uparrow\cdots$). In ferrimagnets ($\uparrow\downarrow\uparrow\cdots$) magnetic moments $M_{\tau,i} = M_\tau$ of all sites i at each sub-lattice τ are the same, but different to a different sub-lattice τ', $M_\tau \neq M_{\tau'}$.
Best known representatives of ferromagnets and antiferromagnets are the elemental $3d$ transition-metals (see also Fig. 2): bcc-Fe, hcp-Co and fcc-Ni are ferromagnetic in the elemental ground state lattice, bcc-Cr and fcc-Mn (high-temperture phase of Mn) are antiferromagnetic. Actually the magnetism of Cr and Mn is much more complicated. Cr shows a frozen sinusoidal spin-density-wave behavior, whose magnetism is characterized by a magnetization, which varies in space like a wave, and which integrated over space, leads to a vanishing total magnetization:

$$\mathbf{m}(\mathbf{r}) = \mathbf{M}\cos\mathbf{q}\mathbf{r} \qquad \int_\infty \mathbf{m}(\mathbf{r})d\mathbf{r} = 0\,. \tag{1}$$

For the ground state of Cr the experimental result $q \approx 0.952\frac{2\pi}{a}$ represents an *incommensurate* spin-density wave as sketched in Fig. 2. The magnetic and structural periodicities are *incommensurately* different. The antiferromagnetism of Cr and its spin-density-wave behavior have recently received considerable new attention, both experimentally [8] and theoretically [9]. In my opinion the frozen sSDW of Cr could not be explained satisfactorily yet on a quantitative level by density functional theory. Obviously, the ferromagnetic state with $\mathbf{q} = (0,0,0)$, and the Nèel antiferromagnetic state are particular examples of the frozen sSDW state. For the "antiferromagnetic approximation" of bcc Cr and γ-Mn, the high-temperature fcc phase of Mn, the wave vector of the magnetization points along one of the cubic axes: $\mathbf{q} = (0,0,q)$ with $q = 2\pi/a$ and positive and negative magnetization are alternating in z direction. Actually "an-

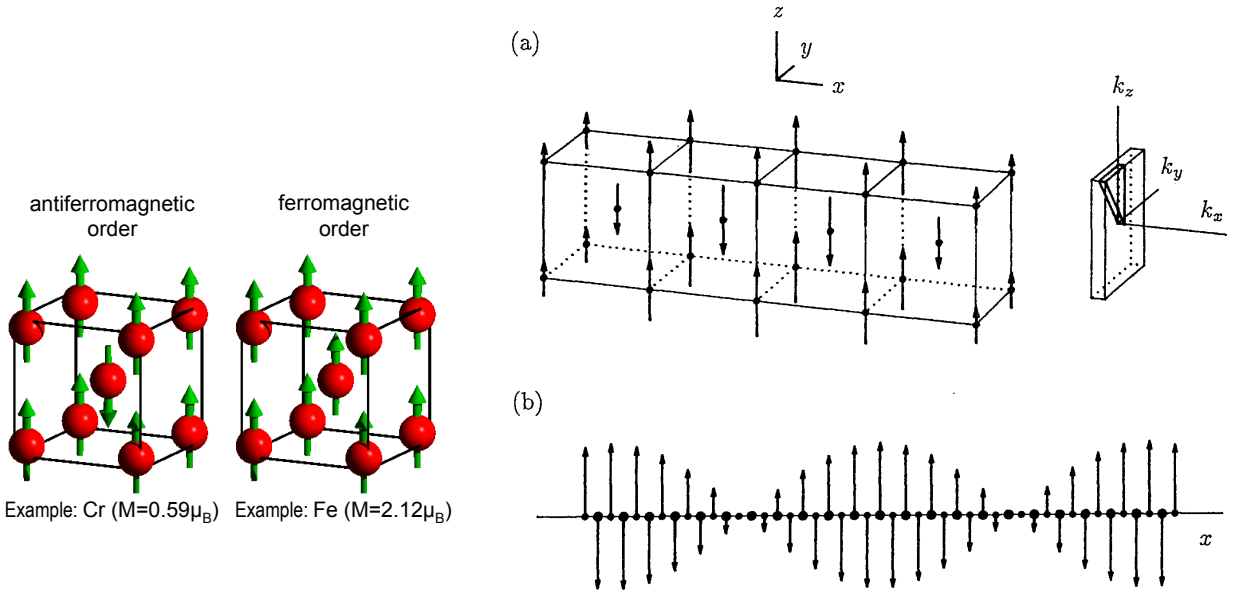


Fig. 2: *Left panel, antiferromagnetic and ferromagnetic structure of bcc Cr and Fe. Right panel, sinusoidal spin density wave in a bcc lattice after Hirai [10]: (a) for* $q = \frac{11}{12}\frac{2\pi}{a}$ *in the region of maximal moments, (b) for* $q = 0.952\frac{2\pi}{a}$*, where the arrows mark the direction and the size of the moments. Small points mark corner atoms, large points central atoms in the bcc unit cell.*

tiferromagnetic" Mn at room temperature has a complicated unit cell containing 58 atoms and will not be considered here any further.
Another well-known example of collinear magnets comes from the field of transition-metal oxides. Transition-metal oxides can be magnetic and they can exhibit a very rich spectrum of magnetic phases. Historic examples include the perovskite compounds $(La_{1-x}Ca_x)MnO_3$ [11]. As function of the Ca concentration x this series of compounds exhibits ferromagnetic (FM) and antiferromagnetic (AFM) properties, which depend upon the relative trivalent and tetravalent manganese ion content. The samples are purely ferromagnetic over a relatively narrow range of composition ($x \sim 0.35$) and show simultaneous occurrence of ferromagnetic and antiferromagnetic phases in the ranges ($0 < x < 0.25$) and ($0.40 < x < 0.5$). Several types of antiferromagnetic structures at $x = 0$ and $x > 0.5$ have been determined. According to the work of Wollan and Koehler [11] one classifies the type of antiferromagnetic phase as AFM-A to AFM-G phase. Fig. 3 gives some examples.
Metallic ferrimagnets are found in compounds containing $4f$ and $3d$ elements. Examples, are the Laves phase compounds, e.g. $GdFe_2$, $Gd_{1-x}Y_{x2}$, $CeFe_2$. A well-known class of examples of insulating ferrimagnets are the ferrites, e.g. $FeO.Fe_2O_3$, $NiO.Fe_2O_3$, $CuO.Fe_2O_3$, ... in the spinel structure. They contain different transition-metal ions (or in case of magnetite Fe_3O_4 it contains two inequivalent iron atoms with two different valencies) carrying two different magnetic moments.

1.2 Non-collinear magnetism

In noncollinear magnetic structures, the magnetization axis $\hat{\mathbf{e}}_M^i$ is not the same for all atoms and in fact it can change direction from site to site. Noncollinear magnetic structures were first

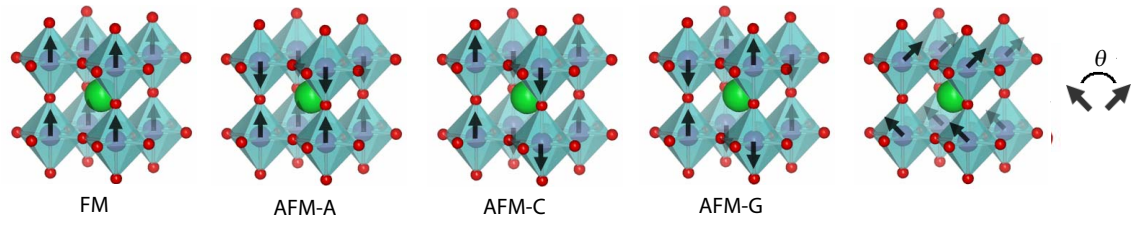


Fig. 3: *Examples of collinear ferromagnetic (FM) and antiferromagnetic (AFM) magnetic states as typical for perovskites, e.g. $(La_{1-x}Sr_x)MnO_3$ (La or Sr, biggest ball green; Mn, mid-size ball blue; O, tiny ball red): FM, AFM-A, AFM-C and AFM-G states, respectively, following the Wollan-Koehler notation [11]. The outer right image represents a non-collinear magnetic state, with a quantization axis, whose direction changes from plane to plane by an angle θ.*

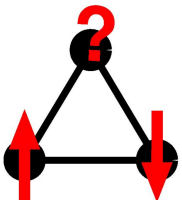


Fig. 4: *Illustration of frustration of AFM interaction between pairs of atoms on triangular lattice.*

discovered experimentally [12, 13] about 50 years ago. Non-collinear magnetism in general and incommensurate spiral spin-density waves in particular are complex magnetic structures which exist in a variety of systems. They often occur in systems with a (i) topologically frustrated antiferromagnetic interaction, e.g. such as for antiferromagnetically interaction magnetic atoms on a triangular lattice, disordered systems, spin-glasses, or (ii) in materials with competing exchange interactions between different neighbors as for example for fcc Fe, the helimagnetism of the lanthanides (see Fig. 5), and multi-component magnets, e.g. $LaMn_2Ge_2$ (see Fig. 12), in exchange bias systems, and molecular magnets, or (iii) on a longer scale in materials with competing interactions of different type, e.g. domain wall. In (i) and (ii) the complex magnetic structure comes purely from exchange and therefore the spiral-states in those helimagnets are often called exchange spirals.

To give the reader at least an example of the richness of the possible magnetic phases I present a series of insulating rare earth (R) perovskites $R\mathrm{MnO_3}$ (R =La, Pr, Nd, ...). $R\mathrm{MnO_3}$ systems crystallize in a perovskite structure with a large $\mathrm{GdFeO_3}$-type distortion [14]. With changing the rare earth atom from La with a large ionic radius to the small ionic radius of Er, the bond angles of Mn-O-Mn decrease and the $\mathrm{GdFeO_3}$-type distortions increase. It was reported on ground of experiments [15], that the magnetic ground state changes from A-type antiferromagnetic (AFM-A) (Fig. 6(a)) to incommensurate magnetic (spiral) state (Fig. 6(d)) to E-type antiferromagnetic (AFM-E) (Fig. 6(c)). The spin-ordering temperatures of $R\mathrm{MnO_3}$ as a function of the in-plane Mn-O-Mn bond angle is shown in Fig. 6.

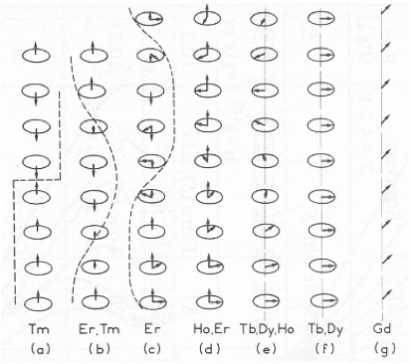


Fig. 5: *Helimagnetism of some $4f$-metals.*

The helimagnetic state of the lanthanides looks like a snapshot of a single magnon at a fixed time. Therefore, these ground states are often called frozen magnons. A magnon is a collective excitation, where the local magnetic moments deviate slightly from the ferromagnetic ground state (or any other magnetic state as ground state) and are typically not stationary states. In order

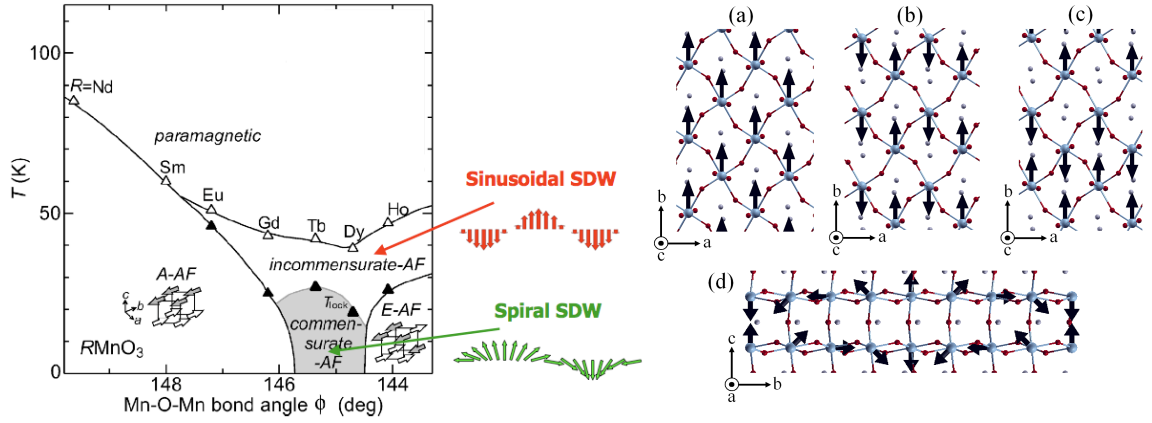


Fig. 6: *Left figure: Spin-ordering temperatures of rare earth (R =Nd, ..., Ho) $RMnO_3$ as a function of the in-plane Mn-O-Mn bond angle according to Kimura et al. [15]. Right figure: Magnetic structures of $RMnO_3$. Black arrows and large gray spheres denote the spin and Mn atoms, respectively. Small gray and red spheres denote the rare earth (R) and O atoms, respectively. All magnetic structures in this figure exhibit an AFM coupling along the c-axis. (a), (b) and (c) are AFM-A, AFM-G and AFM-E, respectively. (d) describes the spiral magnetic structure with a rotation angle of 45° from site to site along the spin-spiral propagation direction.*

to relate the frozen magnon state with the dynamical magnon one relies on an *adiabatic hypothesis*, which conjectures that the slow motion of low-energy magnetic excitations can be decoupled from the fast motion of intersite electron hopping, so that the local electronic structure has time to relax under the constraint that a magnon traverses the system. Under this assumption, the spin-spiral state can therefore be used to simulate the magnon dispersion of a magnetic system in the adiabatic approximation, in particular at very low temperatures, when magnons with long wavelength dominate. At higher temperatures, directional fluctuations of local moments reduce the magnetization. In one-component systems it is well known that the magnetic configuration at each instant shows some degree of short-range order: small regions present almost collinear magnetic moments, with a local spin quantization axis $\hat{\mathbf{e}}_{\mathrm{eloc}}$ not necessarily parallel to the average moment (the global axis $\hat{\mathbf{e}}_{\mathrm{glob}}$). Thus, under the adiabatic assumption, such a thermal fluctuation of a magnetic state can be understood as a snapshot of a thermal fluctuation. This in terns shows the potential that arises if one is able to study the non collinear state one a sophisticated level by a materials dependent theory, such as the spin-dependent density functional theory.

1.3 Exchange mechanisms

In insulating materials the exchange interaction $J_{i,j}$ is typically short-ranged. There exist two categories of types of magnetic interaction:
(i) Direct (potential) exchange: This is driven by minimising potential energy, by reducing wavefunction overlap. Wavefunction overlap is reduced by adding nodes to the wavefunction. This can be done by producing antisymmetric spatial wavefunctions, and so favors symmetric spins, i.e. *ferromagnetic* interactions. This case arises when electrons occupy wavefunctions that overlap in space. In case of wavefunctions localized on different ions the minimization

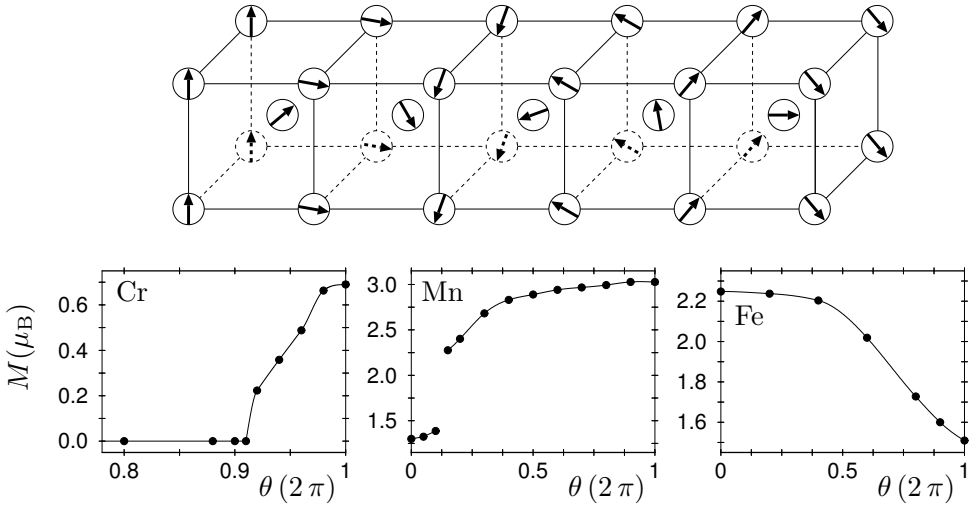


Fig. 7: *Change of the local magnetic moments M in bcc Cr, Mn, and Fe metals as function of the angle θ between the magnetization axes of two consecutive (001) planes. Magnetic moments are aligned ferromagnetically within each (001) plane. Results are calculated for the experimental lattice constants. ($\theta = 2\pi$) $\theta = 0$ corresponds to the (anti-)ferromagnetic state.*

of the Coulomb energy between the different ions lead to symmetric spatial wavefunctions favoring antisymmetric spins, i.e. *antiferromagnetic* interactions.

(ii) Kinetic exchange: This is driven by minimising kinetic energy, by reducing gradients of wavefunctions, i.e. allowing delocalisation of electrons. In insulating compounds with localized electrons the dominating kinetic exchange interaction arises from a process in perturbation calculation involving virtual electron transfer; the resulting interaction is usually *antiferromagnetic*. Kinetic exchange is also sometimes referred to as *superexchange*; we will not use this terminology, as we will use superexchange to refer to a generalisation of this idea, involving hopping via intermediate atoms. This is the historic use of terminology: In this case that transition-metal atoms are that *super* far apart that there is no direct overlap, but still have a strong kinetic exchange. There is an other mechanism favoring ferromagnetism that we can trace back to the discussion of the *double exchange* interaction proposed by Zener [16] for mixed valence manganates in which we have an real exchange of electrons. *Zener's s-d exchange* [17] is a particular mechanism on the basis of kinetic exchange.

There is a potential confusion in the above: these two effects compete, and it is not immediately clear whether in a given situation it is more important to delocalise electrons, or prevent their overlap with other electrons.

(iii) Metallic exchange: In magnetic insulators, the electrons are properly associated with particular atomic sites and the magnetism depends mostly on local or intra–atomic quantities. In metallic materials the exchange interaction is typically long-ranged due to the presence of a Fermi-surface. Magnetism is typically introduced by the $3d$ transition-metal or lanthanide atoms. For example, in $4f$-systems, the exchange interaction between $4f$ localized moments is mediated by $5d$ and $6s$ itinerant electrons and takes the form of a Ruderman-Kittel-Kasuya-Yoshida-type (RKKY) interaction $J(R_j) \propto \cos 2k_F R_j / R_j^3$, for $R_j \ll 1/k_F$, with k_F being the Fermi wave vector, and R_j is the distance of atom j with respect to atom i.

For *itinerant magnetism* or metallic magnetism typical for $3d$-metals, respectively, the electrons responsible for the existence of magnetism hop across the lattice, and thus probe the lattice

and take part in the formation of the Fermi surface. Thus magnetism is intimately interwoven with the electronic structure of the systems. Likewise, the electronic structure is dependent on the atomic arrangement, the composition, and the dimensionality of the system. In transition metals the hopping electrons are correlated, i.e. electrons which hop in and out at a given atomic site share the same orientation of the magnetic spin moment. This finally leads to stable, quasi static local magnetic moments at a given site when averaged over times long compare to hopping times, which are in the order of femto-seconds. This is different to Na, for example, which has fluctuating magnetic moments on the femto-second time scale due to hopping electrons, but no local moment after time averaging. In the spirit of an adiabatic approximation [18, 19] these local magnetic moments may then be treated as classical degrees of freedom relevant to describe pico-second spin dynamics and statistical physics. A typical character of itinerant magnetism is that the magnetic moments are in general not integer, but real numbers, and the value of the moment depends on their relative orientation. This is quantitatively shown in Fig. 7 for bcc Cr, Mn, and Fe.

1.4 Chiral magnetism

Chiral magnets are a particular type of non-collinear magnets that deserve a special mentioning. Systems with lattices lacking inversion symmetry, develop in addition to the above exchange interaction a spin-orbit caused Dzyaloshinskii-Moriya-type [20, 21] antisymmetric exchange interaction, which is one source of origin for multiferroicity in complex oxides and is the origin of the weak ferromagnetism of Fe_2O_3, or of the actinide compounds U_3X_4 (X = P, As or Sb) and U_2Pd_2Sn, to name, but a few popular cases.

Consider a magnetic dimer of two identical magnetic atoms (e.g. blue atoms in Fig. 8) with magnetic moments of unit length, $\mathbf{m} = \hat{\mathbf{e}}$, that are ferromagnetically (↑↑) aligned. If we consider the energy change upon rotating one of the spins, then the energy increase is symmetric with respect of the rotation direction of the moment, say clockwise (↑→), $\mathbf{c} = \hat{\mathbf{e}}_y$, or counterclockwise (↑←), $\mathbf{c} = -\hat{\mathbf{e}}_y$, measured in terms of the vector chirality, $\mathbf{c} = \mathbf{m}_i \times \mathbf{m}_j$, between the magnetic moment $\mathbf{m}_i$ and $\mathbf{m}_j$, if the exchange interaction is caused by the Coulomb interaction. But this is different for the Dzyaloshinskii-Moriya-type (DM) of interaction, which is a chiral interaction, in which one of the two directions is preferred. Thus, the DM-type interaction tries to destabilise a collinear magnetic state (→←) and can introduce a canting of two magnetic moments at different sites (↘↙) or (↗↖), with one of which having a lower energy than that original collinear antiferromagnetic state (→←). Consistent with the periodic lattice of the crystal the ground state for most of the known chiral magnets corresponds to a spiral state. The periodicity of the spiral is given by an interplay between the DM interaction and the Heisenberg exchange. Of course as said in Section 1.2, spiral states may also be supported in non-chiral magnets. In that case, they can result from competing Heisenberg exchange interactions between different neighbors or are due to geometrical frustration and thus are termed exchange spirals. The main feature, which distin-

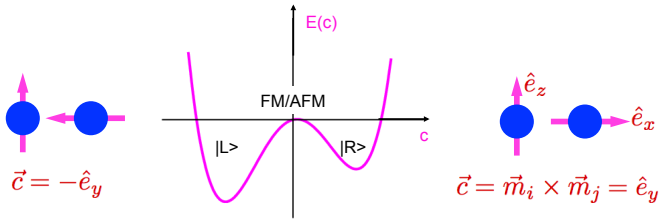


Fig. 8: *Schematic representation of energy versus chirality for a magnetic dimer, represented by two big blue dots. Counterclockwise magnetic structure* $|\mathsf{L}\rangle$ *has lower energy than clockwise rotating structure* $|\mathsf{R}\rangle$.

guishes chiral spirals from exchange ones, is their unique sense of rotation of the magnetization also known as *chirality*.

Chirality is an important property of such magnets. In particular, it supports the stabilization of so-called topologically protected localized magnetic states. Such topologically protected states may occur in chiral magnets as stable or metastable solutions and can be thought of as a two-dimensional analog of the well-known soliton solutions. For the first time, Bogdanov *et al.* [22, 23] has found such solutions in micromagnetic models describing bulk helimagnets, solutions that are now known as chiral magnetic skyrmions. Skyrmions are solutions of smooth non-linear fields [24]. Their topological properties are expressed by a topological number Q, also called winding or skyrmion number, that measures the winding of the magnetic structure when mapped to a unit sphere. For skyrmions, this winding number is a non-zero integer, that means it is not possible to unwind the magnetic structure by external magnetic fields of reasonable strength into a topologically different magnetic structure, e.g. a ferromagnetic state is characterized by a skyrmion number $Q = 0$. Unwinding a skyrmion into a magnetic state with a different topological number goes along with a discontinuous rupture of the smooth skyrmion, which is accompanied by large energy barriers, which effectively makes skyrmions to topologically protected magnetic configurations. Unlike magnetic vortices in nano-discs and magnetic bubbles [25], skyrmions in chiral magnets are stabilized by intrinsic forces, which means that the stabilization of skyrmions does not require particular boundary conditions. Because of the topological protection and localization of the solutions, skyrmions exhibit particle-like properties making them attractive for fundamental research as well as for future applications in various fields, e.g. in *spintronics* [26]. Along this line, a non-zero winding number of a magnetic structure bears some direct physical manifestations in terms of emergent electrical and magnetic fields [27], and electron transport phenomena such as the topological Hall effect [28, 29].

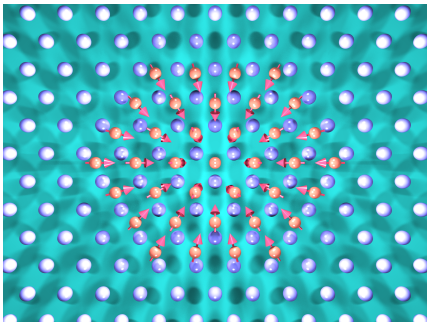


Fig. 9: *Skyrmion with magnetic moments (red arrows) exhibiting a counterclockwise cycloidal rotation locked in a monatomic cluster (red atom with arrows) deposited on a surface (blue-white atoms).*

Interest in chiral magnets surged after recent experiments confirmed the existence of skyrmions in helimagnets using complementary experimental techniques, neutron scattering [30] and Lorentz microscopy [31]. The bulk and thick-film chiral magnets that are currently under intensive scrutiny are the bulk inversion-asymmetric (BIA) cubic helimagnets, e.g. MnSi [30, 32], $Fe_xCo_{1-x}Si$ [31], MnGe [33], FeGe [34, 35], $Mn_{1-x}Fe_xGe$ [36, 37]. There is a track record of experience and knowledge of these materials that goes back to the investigation to the bulk helimagnets studied in the 70's [38].

1.5 Magnetism in low dimensions

Films and nano-clusters on surfaces constitute a rather new field of complex magnetism. (111) oriented substrates provide triangular lattices on which the materials with antiferromagnetic nearest neighbor interaction is frustrated. In low-dimensional systems the coordination number of neighboring atoms is reduced and the balance between on-site Coulomb interaction and kinetic energy of the moving electron is changed towards the on-site exchange strength, which increases in almost all metal films the magnetic moments. Larger magnetic moments support

higher order exchange interactions that lead to much more complex magnetic structures (see for Example 6.2). Then, magnetic films or clusters on surfaces, or sandwiched multilayers or heterostructures experience the presence of an interface. This interface breaks the structure inversion symmetry (SIA) that gives rise to the antisymmetric exchange interaction, whose strength is among other factors determined by the strength of the spin-orbit interaction of the constituent materials.
In this context, recently, a completely new class of chiral magnets has been established, which are made of monolayers and ultrathin films of transition metals (TM), where the DM interaction and the magnetic chirality is introduced exactly by SIA due to the presence of surfaces and interfaces. In some instances the DM interaction introduces chiral ground states, e.g. in Mn/W(110) [1], Mn/W(100) [39], Fe/Ir(111) [40, 41, 42], Pd/Fe/Ir(111) [43], bi-atomic Fe chains on the (5×1)-Ir(001) surface [44] in other instances chiral domain walls, e.g. in Fe/W(110) [45, 46], Ni/Fe/Cu(001) [47] or the chirality alters dynamical properties as for the magnon dispersion in Fe/W(110) [48].

2 Spin Models

2.1 Heisenberg model

Assuming strong (size is independent of relative orientation to other moments) local magnetic moment formation e.g. at $3d$ transition-metal sites through intraatomic exchange (Hund's $1^{\rm st}$ rule), we can now attempt to predict how the local spins cooperate on a global length scale at finite temperature, to explain the magnetic ground state of magnetic transition-metal systems or how they respond to excitations. This, however, can be a highly nontrivial task. In cases, for example, where competing exchange interactions between neighboring transition-metal atoms cannot be satisfied, exchange interactions are frustrated giving rise to a multitude of possible spin-structures depending on slight modifications of the exchange interactions. A trivial example of frustration is realized in an one-dimensional Heisenberg chain of nearest neighbor ferromagnetic interaction $J_1 > 0$ and antiferromagnetic next-nearest neighbor interaction $J_2 < 0$, with $E = -\sum_n [J_1 \mathbf{S}_n \cdot \mathbf{S}_{n+1} + J_2 \mathbf{S}_n \cdot \mathbf{S}_{n+2}]$. If $|J_2| > J_1/4$ the ferromagnetic state becomes unstable and a spin-spiral state with a wave vector $\cos q = J_1/(4|J_2|)$ has the lowest energy. This model explains for example the helimagnetic state of the lanthanides displayed in Fig. 5. For more information on the spin-spiral state see Eq. (45). Since the work of Heisenberg [49], the magnetism of complex spin structures of magnetic insulators has been almost exclusively discussed within the framework of model Hamiltonians, first and foremost the Heisenberg Hamiltonian, to which we refer in its classical version, replacing the Pauli matrices $(\sigma_x, \sigma_y, \sigma_z)$ with the $SU(2)$ symmetry by classical vectors S_x, S_y, S_z of $O(3)$ symmetry, neglecting the possibility of quantum fluctuations:

$$\mathcal{H}_{\text{2-spin}} = -\sum_{i,j} J_{ij}\, \mathbf{S}_i \cdot \mathbf{S}_j \qquad \text{and} \qquad \mathbf{S}_i^2 = S^2, \text{ for all } i\ , \tag{2}$$

i.e. the spins of the d-electrons localized on the lattice sites i, j are considered as classical vectors $\mathbf{S}$, with the assumption that the spins on all lattice sites have the same magnitude S. The exchange interaction between the spins is isotropic and described by the pair interaction J_{ij}. In localized model spin systems the J_{ij} can be safely approximated by the ferromagnetic ($J_1 > 0$) or antiferromagnetic ($J_1 < 0$) nearest-neighbor (n.n.) interaction, i.e. $J_{ij} = 0$ for all

i, j, except for $J_{n.n.} = J_1$. Although in magnetic insulators J_1 often dominates over the rest of the further distant pairs, an attempt to reproduce the magnetic complexity or the T_c solely from J_1 produces results of limited quality or often simply fails. In magnetic insulators the inclusion of several neighbors is important. Please notice, in the literature one finds a set of slightly different definitions of the Heisenberg model. I have chosen here a negative sign. Further, I have not corrected the double counting of sites $(i,\, j)$ and $(j,\, i)$ by a prefactor $1/2$.
In the Lecture B1 *Model Hamiltonians and Second Quantization* we have seen that the Heisenberg model, which is an effective spin-model, can be derived from the Hubbard model which describes the motion of electrons on the lattice at the presence of on-site electron correlation. Analogously, exchange interactions beyond the Heisenberg model can be motivated from a perturbation expansion of the Hubbard model [50] in powers of t^2/U in the limit of large U (see Eq. 32 of Lecture B1). Expanding the Hubbard model into a spin model, replacing the spin operators by classical spin vectors, a second order perturbation expansion reproduces the classical Heisenberg model with $J_1 = 2t^2/U$. J arises from the same kind of kinetic exchange process kinetic exchange process that we learned about in connection with the H_2 problem.
For itinerant transition-metal magnets and depending on the Fermi surfaces J_{ij} between distant pairs are required to model the properties with reliability. Distances up to the 10th nearest neighbor are not uncommon. In most cases also in transition-metal magnets the interaction between the 1^{st} neighbor is the largest.

2.2 Higher-order exchange

The fourth order perturbation treatment (the third order is zero in the absence of spin-orbit interaction) yields two additional terms of different form. One is the four-spin exchange interaction (4-spin):

$$\mathcal{H}_{\text{4-spin}} = \sum_{ijkl} K_{ijkl}\big[(\mathbf{S}_i\mathbf{S}_j)(\mathbf{S}_k\mathbf{S}_l) + (\mathbf{S}_j\mathbf{S}_k)(\mathbf{S}_l\mathbf{S}_i) - (\mathbf{S}_i\mathbf{S}_k)(\mathbf{S}_j\mathbf{S}_l)\big]\ . \tag{3}$$

The 4-spin interaction is so-called ring-exchange and arises from four consecutive hops of electrons from one spin configuration to the spin-flipped one. For the sake of simplicity let us consider a four-site ring with nearest neighbor hopping. Denoting the eigenstates as $|S_1^z S_2^z S_3^z S_4^z\rangle$, etc., we find for example

$$|\uparrow\downarrow\uparrow\downarrow\rangle \rightsquigarrow |0(\uparrow\downarrow)\uparrow\downarrow\rangle \rightsquigarrow |\downarrow\uparrow\uparrow\downarrow\rangle \rightsquigarrow |\downarrow\uparrow(\uparrow\downarrow)0\rangle \rightsquigarrow |\downarrow\uparrow\downarrow\uparrow\rangle\,. \tag{4}$$

Such a change is effected by $S_1^- S_2^+ S_3^- S_4^+$ (with e.g. $S_3^-|\uparrow\uparrow\uparrow\uparrow\rangle = |\uparrow\uparrow\downarrow\uparrow\rangle$). Since the total Hamiltonian is spin-rotationally invariant, this must be part of the isotropic term $(\mathbf{S}_1\mathbf{S}_2)(\mathbf{S}_3\mathbf{S}_4)$. By symmetry, the term $(\mathbf{S}_2\mathbf{S}_3)(\mathbf{S}_4\mathbf{S}_1)$ must have the same coefficient. Finally, we can convince ourselves that there must exist a term like $(\mathbf{S}_1\mathbf{S}_3)(\mathbf{S}_2\mathbf{S}_4)$ as well. An algebraic trick to facilitate the spin exchange between two sites i and j is to introduce the permutation operator (also known as Dirac spin-exchange operator [51], $\mathcal{P}_{ij} = 2(\mathbf{S}_i \cdot \mathbf{S}_j + 1/2)$ with $\mathcal{P}_{12}|\sigma_1\sigma_2\rangle = |\sigma_2\sigma_1\rangle$. Enumerating all the processes, one finds the four-site term $\mathcal{H}_{\text{4-spin}} = \sum_{ijkl} \mathcal{P}_{ij}\mathcal{P}_{jk}\mathcal{P}_{kl}\mathcal{P}_{li}$ (or plaquette exchange) [50] given in Eq. (3) with $K_{\text{n.n.}} = 5t^4/U^3$. This is strictly true only for systems with localized $S = 1/2$ spins. For $S > 1/2$ magnets (which may be described by degenerate Hubbard models), more possibilities have to be considered and the first non-trivial term, which appears is the biquadratic exchange:

$$\mathcal{H}_{\text{biquadr}} = \sum_{ij} B_{ij}(\mathbf{S}_i \cdot \mathbf{S}_j)^2\ . \tag{5}$$

The exchange parameters J_{ij}, K_{ijkl}, and B_{ij} depend on the details of the electronic structure, e.g. the Fermi surface in case of metals. Since the higher-order terms scale with the magnetic moment to 4th-order, their relative importance increases with increasing magnetic moments. They should therefore become important for Fe atoms and for transition-metals atoms at surfaces. Although the higher-order magnetic interactions are in general small may be 5% of J, their value lies in the fact that they can lift the degeneracy of those magnetic states that are degenerate in the Heisenberg model. Further, there is always a chance that they play a role in frustrated magnets, where the competition between leading-order terms leaves the questions of order for the weaker terms to decide [52, 53].

2.3 Dipole-dipole interaction

For completeness I would like to mention that a piece of magnetic material is typically magnetically anisotropic. This means that the total energy depends on the orientation of the magnetization as measured with respect to the crystal axes and the sample shape. Thus, besides the isotropic exchange interaction there are additional interactions, so-called anisotropic interactions. This orientation dependent energy contribution is called the *magnetic anisotropy energy* (MAE), $E_{\rm MAE}$. Without this effect of the magnetic anisotropy, magnetism would have been hard to discover and possibly useless. In some way or the other, almost all applications of magnetic materials hinge on the fact that it is easier to magnetize a magnetic material in one direction than another. The magnetic anisotropy is responsible for the occurrence of easy and hard axes. The magnetic anisotropy has two origins:
(i) one is the classical dipole-dipole interaction between magnetic moments of localized spins $\mathbf{M}$ at transition-metal sites.

$$\mathcal{H}_{\rm dip} = \sum_{i\neq j} \frac{1}{r_{ij}^3}(\mathbf{M}_i \cdot \mathbf{M}_j - 3(\mathbf{M}_i \cdot \hat{\mathbf{r}}_{ij})(\mathbf{M}_j \cdot \hat{\mathbf{r}}_{ij}))\,, \tag{6}$$

where $\hat{\mathbf{r}}_{ij}$ is the unit vector pointing in the direction of $\mathbf{r}_j - \mathbf{r}_i = \mathbf{r}_{ij}$. It is worth mentioning that $\mathbf{M}$ is the total magnetic moment of the ion including the orbital magnetic contribution. Thus, the effect of the dipole-dipole interaction will be relatively strong for rare-earth ions ($\mathbf{M} \propto \mathbf{J} = \mathbf{L} + \mathbf{S}$), while for $3d$ transition-metal magnets and transition-metal oxides, $\mathbf{M} \propto \mathbf{S}$. For antiferromagnets the contribution is zero or very small (in case of weak ferromagnetism due to e.g. Dzyaloshinskii-Moriya interaction). Due to the long range nature of the dipole-dipole interaction it contributes mostly to the *shape anisotropy* of the MAE.

2.4 Anisotropy of on-site interaction

(ii) The second contribution arises from the spin-orbit coupling,

$$\mathcal{H}_{\rm SO} = \xi\, \mathbf{L} \cdot \mathbf{S}\,, \tag{7}$$

leaving a degeneracy $(2|\mathbf{J}| + 1)$, which relates the spin-degree of freedom to the motion of the electrons in the lattice and contributes exclusively to the *magneto-crystalline anisotropy* (MCA). Its strength depends crucially on the symmetry of the lattice and the orbital moment in the crystal. It is frequently called the *single-ion anisotropy*.
In crystal field theory we learn that the crystal field Hamiltonian is described by a real potential, enabling the choice angular wavefunctions usinging only real numbers. iThus one may choose

eigenstates which are also real. Since the angular momentum is a purely imaginary operator, it implies that for such an eigenstate:

$$\langle\psi|\mathbf{L}|\psi\rangle = 0\,, \tag{8}$$

i.e. angular momentum is quenched, and only the spin degree of freedom remains. There is however a caveat if the crystal field does not entirely remove the degeneracies, then from a pair of degenerate real wavefunctions, one may construct combinations for which the above equality does not hold. In magnetic insulators it is usually such that degeneracies do not survive, since it becomes energetically more favorable to pay the (quadratic) elastic cost of deforming a symmetric lattice to gain a linear decrease in electronic energy from splitting a degenerate state. Assuming complete quenching of angular momentum, i.e. $\langle\psi|\mathbf{L}|\psi\rangle = 0$, it is clear that the spin-orbit perturbation at site i, $\mathcal{H}_{\mathrm{SO}}(i) = \xi\,\mathbf{L}_i \cdot \mathbf{S}_i$ has no effect at first order in perturbation theory, as there is no remaining degeneracy to break. However, this term may still have an effect in second (and higher) order perturbation theory. Considering an unknown spin state, but integrating out orbital excitations perturbatively, one may write:

$$\mathcal{H}^{(2)}_{\mathrm{MCA}} = \delta H = |\xi|^2 \sum_n \frac{\langle 0|L_\mu|n\rangle\langle n|L_\nu|0\rangle}{E_0 - E_n} S_{\mu\nu} = -S_\mu \Lambda_{\mu\nu} S_\nu\,, \tag{9}$$

where μ, ν run over the three cartesian components x, y, z and generally $\Lambda_{\mu\nu}$ is not proportional to the identity matrix, but depends on the ordering of the excited orbitals. Rotating the real, symmetric 3×3 tensor into the principle axes Λ is conventionally written as $\Lambda_{\mu\mu} = K_{2\mu}\delta_{\mu\mu}$, with $K_2 \propto |\xi|^2$ and the uniaxial anisotropy at all sites i is summarized to

$$\mathcal{H}^{(2)}_{\mathrm{MCA}} = \sum_i K_{2;ix} S^2_{i,x} + K_{2;iy} S^2_{i,y} + K_{2;iz} S^2_{i,z}. \tag{10}$$

This equation holds for transition-metal atoms with a two fold symmetry in all three directions. This is a rare case. Frequently, transition-metal ions are located in a cubic or octahedral environment. Then all terms in second order perturbation theory are zero and the first nonvanishing contribution result from the 4th order perturbation theory. The magnetic anisotropy changes to

$$\mathcal{H}^{(4)}_{\mathrm{MCA}} = \sum_i K_{4;i}\,(S^4_{i,x} + S^4_{i;y} + S^4_{i,z})\,, \tag{11}$$

with an anisotropy strength $K_4 \propto |\xi|^4$. This contribution is small. Frequently, the system has a Jahn-Teller distortion, that lowers the symmetry and then a second order contribution is added with an additional smallness parameter, the displacement of the atom that breaks the symmetry.

2.5 Antisymmetric exchange: Dzyaloshinskii-Moriya interaction

In 1957 Dzyaloshinskii [20] predicted on the basis of symmetry arguments the presence of a unidirectional magnetic interaction of the form:

$$\mathcal{H}_{\mathrm{DM}} = \sum_{i,j} \mathbf{D}_{ij} \cdot (\mathbf{S}_i \times \mathbf{S}_j)\ , \tag{12}$$

where the Dzyaloshinskii-vector, $\mathbf{D}$, is a vector, which depends on the symmetry of the system and on the real space direction given by two sites i and j. The presence of the DM interaction has far-reaching consequences. Depending on the sign, the symmetry properties, and the

value of the Dzyaloshinskii vector $\mathbf{D}_{ij}$, collinear uniaxial ferro- or antiferromagnetic structures are becoming unstable and are instead replaced by a directional noncollinear magnetic structure of one specific chirality, $\mathbf{c} = \mathbf{S}_i \times \mathbf{S}_{i+1}$, either a right-handed ($c > 0$) or left-handed ($c > 0$) one. The interaction is fundamentally different to the direct, kinetic, or metallic exchange interactions we have been discussed before that are caused exclusively by the Coulomb interaction. Those are symmetric, i.e. two magnetic configurations with right-handed ($\nwarrow\nearrow$) or left-handed ($\nearrow\nwarrow$) alignment of the magnetic moments have the same energy. The nature of the DM interaction prefers non-collinear magnetic structures and is responsible for weak ferromagnetism, i.e. the occurrence of small net magnetic moments of otherwise antiferromagnetic materials. In this context one may think of hematite (Fe_2O_3), a mineral with an antiferromagnet exchange ($J > 0$) as an example. The canting angle can be estimated from a simple two site model

$$\mathcal{H}_{\rm DM} = -J\mathbf{S}_1 \cdot \mathbf{S}_2 \; + \; \mathbf{D} \cdot (\mathbf{S}_1 \times \mathbf{S}_2) \; .$$

Assuming that the spins $\mathbf{S}_1$ and $\mathbf{S}_2$ are placed in the same plane, then the energy depends only on the angle θ between the spins ($\searrow\swarrow$)

$$E = -JS^2 \cos\theta + DS^2 \sin\theta \; ,$$

with the canting angle $\tan(\pi - \theta_{\rm min}) = -D/J$. $\mathcal{H}_{\rm DM}$ has the most general form of an antisymmetric tensor of 2$^{\rm nd}$ rank which is second order in spin $\mathbf{S}$. Typically, the combination of hopping of spin-polarised electrons and the presence of spin-orbit interaction can produce such a term. The microscopic details depend on the solid and are described in different microscopic models. In my opinion the minimal model to obtain an antisymmetric exchange interaction consists of 3 atoms, whereby at least two of which are magnetic and and one of which carries spin-orbit interaction. The most famous one is the model of Moriya explaining the DM interaction in magnetic insulators. He derived also rules under which circumstances the DM interaction will occur. These rules are general and should apply also to metallic systems.

Magnetic Insulators: Moriya [21] explained the origin of the interaction for transition-metal oxides on the basis of a microscopic picture, which combines the kinetic (super-)exchange due to hopping electrons that would be described by the Heisenberg model, with the *on-site* spin-orbit coupling at each site that we have discussed in Section 2.4, giving rise to virtual excitations between the ground state g and excited states n. That is, why this interaction carries both names today and is called the *Dzyaloshinskii-Moriya interaction*. It is also referred to as *antisymmetric exchange* or *antisymmetric superexchange*. We discuss this here on the basis of a small two-site model and consider:

$$H = \xi\, \mathbf{L}_1 \cdot \mathbf{S}_1 \; + \; \xi\, \mathbf{L}_2 \cdot \mathbf{S}_2 \; + \; \hat{J}\mathbf{S}_1 \cdot \mathbf{S}_2 \tag{13}$$

where $\hat{J}$ is an operator that depends on what orbital state the electron is in; at the ground state g or at one of the excited states n. Normally this fact is irrelevant because the orbital states are split, and so one may normally consider $\hat{J} \to J = \langle 0|J|0\rangle$, but combined with the spin-orbit coupling we may write consistent with the schematic picture of the hopping paths of the electrons given in Fig. 10:

$$\mathcal{H}_{\rm DM} = \delta H = -\xi \sum_{n_1} \left[\frac{\langle g_1 g_2|\mathbf{L}_1 \cdot \mathbf{S}_1|n_1 g_2\rangle\langle n_1 g_2|\hat{J}\mathbf{S}_1 \cdot \mathbf{S}_2|g_1 g_2\rangle}{E_{n_1 g_2} - E_{g_1 g_2}} \right. \tag{14}$$

$$\left. + \; \frac{\langle g_1 g_2|\hat{J}\mathbf{S}_1 \cdot \mathbf{S}_2|n_1 g_2\rangle\langle n_1 g_2|\mathbf{L}_1 \cdot \mathbf{S}_1|g_1 g_2\rangle}{E_{n_1 g_2} - E_{g_1 g_2}} \right] + (1 \Leftrightarrow 2).$$

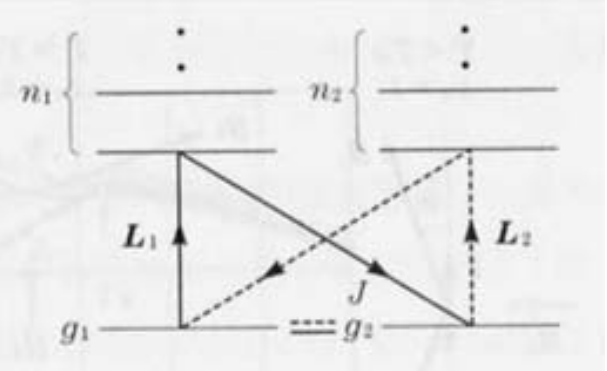


Fig. 10: *Interaction channels of hopping electrons that give rise to the Dzyaloshinskii-Moriya interaction.*

This may be simplified because, for real wavefunctions one has $\langle g_1|\mathbf{L}_1|n_1\rangle = \langle n_1|\mathbf{L}_1|g_1\rangle^* = -\langle n_1|\mathbf{L}_1|g_1\rangle$ and $\langle n_1 g_2|\hat{J}|g_1 g_2\rangle = J(n_1 g_2|g_1 g_2) = J(g_1 g_2|n_1 g_2) = \langle g_1 g_2|\hat{J}|n_1 g_2\rangle$, hence

$$\mathcal{H}_{\rm DM} = -2\xi \sum_{n1} \frac{\langle g_1|\mathbf{L}_1|n_1\rangle\langle n_1 g_2|\hat{J}|g_1 g_2\rangle}{E_{n_1 g_2} - E_{g_1 g_2}} [\mathbf{S}_1, \mathbf{S}_1 \cdot \mathbf{S}_2] \; + \; (1 \Leftrightarrow 2). \tag{15}$$

Writing the sum over n_1 as D_1, and using the commutation relations to write, $[\mathbf{S}_1, \mathbf{S}_1 \cdot \mathbf{S}_2] = -i\mathbf{S}_1 \times \mathbf{S}_2$ and $[\mathbf{S}_2, \mathbf{S}_1 \cdot \mathbf{S}_2] = i\mathbf{S}_2 \times \mathbf{S}_1$ one can write

$$\mathcal{H}_{\rm DM} = (\mathbf{D}_1 - \mathbf{D}_2) \cdot (\mathbf{S}_1 \times \mathbf{S}_2) \;, \quad \text{with} \quad \mathbf{D}_1 = -2i\,\xi \sum_{n1} \frac{\langle g_1|\mathbf{L}_1|n_1\rangle}{E_{n_1 g_2} - E_{g_1 g_2}} J(n_1 g_2|g_1 g_2) \tag{16}$$

The strength of the Dzyaloshinskii-Moriya interaction, i.e. the length of the Dzyaloshinskii-vector, respectively, can be estimated to $D \approx |\frac{\xi}{\Delta E}|J$, and is thus first order in the spin-orbit interaction, in difference to the single-ion anisotropy (Section 2.4) that is at least second order. It is thus clear, that D is equal zero, $D = 0$, in case of a mirror symmetry or point of inversion between the two atoms. In other words, the DM interaction exists only between two *different* atoms. However, *different* need not mean different species of ion, but can just mean different chemical environments, such as in a bipartite lattice, where the symmetry of orbitals may be different between the two types of lattice site. In article [21], Moriya derived the following symmetry rules (also expressed schematically in Fig. 11: Considering two ions located at sites A and B, respectively, and the point bisecting the straight line AB is denoted by C. The rules are:

(i) When a center of inversion is located at C: $\mathbf{D} = 0$.

(ii) When a mirror plane perpendicular to AB passes through C: $\begin{cases} & \mathbf{D} \parallel \text{mirror plane} \\ \text{or} & \mathbf{D} \perp AB. \end{cases}$

(iii) When there is a mirror plane including A and B: $\mathbf{D} \perp$ mirror plane.

(iv) When a two-fold rotation axis perpendicular to AB passes through C: $\mathbf{D} \perp$ two-fold axis

(v) When there is an n-fold ($n \geq 2$) along AB: $\mathbf{D} \parallel AB$

The DM interaction plays an important role in the field of multiferroics. It is one of the possible interactions that causes a linear magnetoelectric effect, $\Phi = -\alpha_{ij} E_i B_J$, with the electric polarization $P_i = -\partial\Phi/\partial E_i = \alpha_{ij} B_j$ and the magnetization $M_i = -\partial\Phi/\partial B_i = \alpha_{ji} E_j$ coupling.

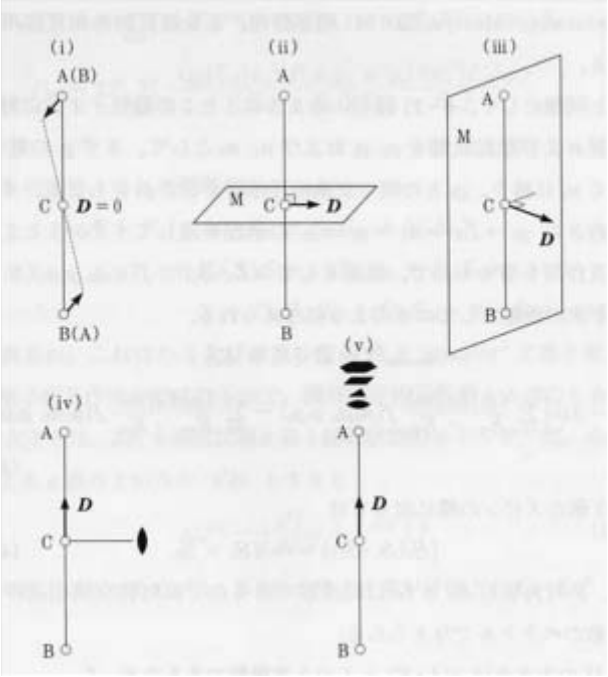


Fig. 11: *Direction of the Dzyaloshinskii-Moriya vector for a diatomic molecules located at sites A and B according to the symmetry rules of Moriya [21].*

An example is Cr_2O_3. The DM interaction causing magnetic structures with a unique rotational order – we think in particular at cycloidal spirals $\mathbf{M} = M(\hat{\mathbf{x}}\cos\mathbf{qR} + \hat{\mathbf{y}}\sin\mathbf{qR})$, $\mathbf{R} \perp (\hat{\mathbf{x}}\hat{\mathbf{y}})$ – can also be a source of inversion symmetry breaking in a structurally symmetric compound being the origin of magnetic ferroelectrics also known as improper multiferroelectrics [54]. In such spirals ferroelectric polarization P is introduced that is proportional to the outer product of $\mathbf{P} \propto [\hat{\mathbf{z}} \times \mathbf{q}_x]$. In these systems the magnetoelectric coupling is rather strong. This new class of materials include for example $TbMnO_3$, $TbMn_2O_5$, $Ni_3V_2O_8$, $MnWO_4$, $CoCr_2O_4$, $CuFeO_2$ etc..

Magnetic Metals: The antisymmetric spin-interaction in metals had been discussed in different microscopic models that go back to Smith [55], Fert and Levy [56, 57] and the Kanamori-group [58]. Fert and Levy [56] derived an expression for this anisotropic exchange interaction of two magnetic atoms in spin-glasses doped with heavy impurity atoms, i.e. impurity atoms with a large nuclear number and thus a large spin-orbit interaction, which is of the form

$$\mathcal{H}_{\rm DM} = -V(\xi)\frac{\sin\left[k_{\rm F}(R_A + R_B + R_{AB}) + \eta\right]\hat{\mathbf{R}}_A \cdot \hat{\mathbf{R}}_B}{R_A R_B R_{AB}}(\hat{\mathbf{R}}_A \times \hat{\mathbf{R}}_B)(\mathbf{S}_A \times \mathbf{S}_B) \qquad (17)$$

where $\mathbf{R}_A = R_A\hat{\mathbf{R}}_A$ and $\mathbf{R}_B = R_B\hat{\mathbf{R}}_B$ are the positions of the magnetic atoms measured from the nonmagnetic impurity and R_{AB} is the distance between the atoms A and B. $V(\xi)$ is a term that depends of the spin-orbit coupling constant of the nonmagnetic atom, ξ, $k_{\rm F}$ is the Fermi vector and η the phase shift induced by the impurity. The sinus term reflects the RKKY-type character of the interaction, while the two cross products determine the symmetry of the interaction. How far this model holds for real materials has not thoroughly been investigated yet on the basis of density functional theory. The model should be also applicable to adatoms on surfaces. Here, Rashba electrons, electrons in a surface state that is subject to the presence

of spin-orbit interaction as well as the structure inversion asymmetry due to the presence take the role of the heavy atoms and produce also an antisymmetric exchange interaction, which has recently been investigated by Bouaziz *et al.* [59], which is proportional to the Rashba strength and falls off with a slower power law that is $1/R^2$ or $1/R$ for large distances.
At the end we summarise all spin-spin interactions, due to exchange with and without spin-orbit interaction, that can be described up to second order by the Hamiltonian

$$H = \sum_i K(\mathbf{S}_i) + \sum_{ij(i\neq j)} \mathbf{S}_i \mathsf{J}_{ij} \mathbf{S}_j \tag{18}$$

Here, the first term describes the on-site anisotropy energy and the second term (where $i \neq j$) includes the intersite exchange interactions and the Dzyaloshinskii-Moriya interaction. The quantities J_{ij} are 3×3 matrices (second rank tensors in real space). They can be decomposed in symmetric and antisymmetric parts as: $\mathsf{J}_{ij} = J_{ij}1 + \mathsf{J}^{\mathrm{S}}_{ij} + \mathsf{J}^{\mathrm{A}}_{ij}$ with J_{ij} being the usual Heisenberg exchange interaction $J_{ij} = 1/3\,\mathrm{Tr}\mathsf{J}_{ij}$, $\mathsf{J}^{\mathrm{S}}_{ij}$ the *anisotropic exchange interaction* or sometimes called two-ion anisotropy which is the (traceless) symmetric part $\mathsf{J}^{\mathrm{S}}_{ij} = \frac{1}{2}\left(\mathsf{J}_{ij} + \mathsf{J}^{t}_{ij}\right) - J_{ij}1$, and $\mathsf{J}^{\mathrm{A}}_{ij}$ the antisymmetric part: $\mathsf{J}^{\mathrm{A}}_{ij} = \frac{1}{2}\left(\mathsf{J}_{ij} - \mathsf{J}^{t}_{ij}\right)$. The latter can be brought in connection with the Dzyaloshinskii-Moriya interaction constants via the antisymmetric tensor, $D^x_{ij} = \varepsilon_{xyz} J^{\mathrm{A}yz}_{ij}$, and the spin-spin interaction in its general form is then written in terms of the Dzyaloshinskii-Moriya vectors as

$$\mathbf{S}_i \mathsf{J}_{ij} \mathbf{S}_j = J_{ij}\mathbf{S}_i\mathbf{S}_j + \mathbf{S}_i \mathsf{J}^{\mathrm{S}}_{ij} \mathbf{S}_j + \mathbf{D}_{ij}(\mathbf{S}_i \times \mathbf{S}_j). \tag{19}$$

Above not much has been said about the anisotropic exchange interaction. It is second order in spin-orbit interaction, while the Dzyaloshinskii-Moriya interaction is first order. It is important to notice that all these exchange interactions and those beyond the simple models discussed here are included (in many case to high accuracy) in density functional theory, which we describe next, although there one-to-one correspondence it not always trivial.

3 Spin Density Functional Theory

In 1964 Hohenberg and Kohn [60] worked out two central theorems that form the basis of density functional theory (DFT): For a system of N particles (e.g. electrons) moving in an external potential $v(\mathbf{r})$ (caused by e.g. nuclei) in a non-degenerate ground state (i) the many-body wavefunction Ψ and $v(\mathbf{r})$ are uniquely determined by the particle density distribution $n(\mathbf{r})$ and (ii) there exists an energy functional of this density, $E[n(\mathbf{r})]$, which is stationary with respect to variations of the ground-state density. These two theorems allow – at least in principle – the determination of the ground-state density and energy of a N-particle system. Extracting the classical Coulomb interaction energy, such a Hohenberg-Kohn energy functional takes the form

$$E[n(\mathbf{r})] = \int v(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + \frac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + G[n(\mathbf{r})] \tag{20}$$

where the functional $G[n(\mathbf{r})]$ has to be approximated.
In the Kohn-Sham theory [61], the kinetic energy T_0 of a non-interacting electron gas in its ground state with a density distribution $n(\mathbf{r})$ is further extracted from $G[n(\mathbf{r})]$, so that a new functional

$$E_{\mathrm{xc}}[n(\mathbf{r})] = G[n(\mathbf{r})] - T_0[n(\mathbf{r})] \tag{21}$$

remains to be determined. $E_{\rm xc}$ is now called exchange-correlation energy functional, since without $E_{\rm xc}$ our energy functional E would yield just the Hartree energy. If we take into account that particle conservation, i.e. $N = \int n(\mathbf{r})d\mathbf{r}$ has to be ensured, we can formulate the stationarity of E in Eq. (20) with respect to variations of the ground-state density as

$$v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \frac{\delta T_0}{\delta n(\mathbf{r})} + \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})} - \lambda = 0 \tag{22}$$

where the Lagrange parameter λ ensures the particle conservation. Expressing the kinetic energy of the non-interacting particles via their wavefunctions, ϕ_i, we can recast Eq. (22) in the form of an effective single particle Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + \frac{\delta E_{\rm xc}}{\delta n(\mathbf{r})}\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{23}$$

which has to be solved self-consistently since $n(\mathbf{r}) = \sum_{i=1}^{N}|\phi_i(\mathbf{r})|^2$. The last term of the Hamiltonian is called the exchange-correlation potential. For more detailed information see Lecture A2 *Introduction to Density Functional Theory*.

3.1 The density and potential matrix

In 1972 von Barth and Hedin extended this concept to spin-polarized systems [62], replacing the scalar density by a hermitian 2×2 matrix $\mathsf{n}(\mathbf{r})$. If $\psi_\alpha(\mathbf{r})$ is the field operator for a particle of spin α, a component of the spin-density matrix can be defined as

$$n_{\alpha\beta}(\mathbf{r}) = \langle\Psi|\psi^+_\beta(\mathbf{r})\psi_\alpha(\mathbf{r})|\Psi\rangle. \tag{24}$$

The potential matrix corresponding to this spin-density matrix is denoted as $\mathsf{v}(\mathbf{r})$ and replaces the scalar potential. Then, we can write Eq. (23) in the form

$$\left[\left(-\frac{\hbar^2}{2m}\nabla^2 + \sum_\alpha\int\frac{n_{\alpha\alpha}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}'\right)\mathsf{I} + \mathsf{v}(\mathbf{r}) + \frac{\delta E_{\rm xc}}{\delta\mathsf{n}(\mathbf{r})}\right]\begin{pmatrix}\phi_i^{(+)}(\mathbf{r})\\ \phi_i^{(-)}(\mathbf{r})\end{pmatrix} = \epsilon_i\begin{pmatrix}\phi_i^{(+)}(\mathbf{r})\\ \phi_i^{(-)}(\mathbf{r})\end{pmatrix} \tag{25}$$

where I is a 2×2 unit matrix and the exchange-correlation potential, v, is now also a 2×2 matrix. In terms of the Kohn-Sham wavefunctions, the density matrix can now be written as

$$n_{\alpha\beta}(\mathbf{r}) = \sum_{i=1}^{N}\phi_i^{*\alpha}(\mathbf{r})\phi_i^{\beta}(\mathbf{r}) \quad \text{where} \quad \alpha,\beta = (+),(-). \tag{26}$$

Using the Pauli matrices, $\boldsymbol{\sigma}$, the density matrix can be decomposed in a scalar and a vectorial part, corresponding to the charge and magnetization density:

$$\mathsf{n}(\mathbf{r}) = \frac{1}{2}\left(n(\mathbf{r})\mathsf{I} + \boldsymbol{\sigma}\cdot\mathbf{m}(\mathbf{r})\right) = \frac{1}{2}\begin{pmatrix} n(\mathbf{r})+m_z(\mathbf{r}) & m_x(\mathbf{r})-im_y(\mathbf{r})\\ m_x(\mathbf{r})+im_y(\mathbf{r}) & n(\mathbf{r})-m_z(\mathbf{r})\end{pmatrix}. \tag{27}$$

Likewise, the potential matrices can be written in terms of a scalar potential and magnetic field, $\mathbf{B}(\mathbf{r})$:

$$\mathsf{v}(\mathbf{r}) = v(\mathbf{r})\mathsf{I} + \mu_B\boldsymbol{\sigma}\cdot\mathbf{B}(\mathbf{r}) \quad \text{and} \quad \mathsf{v}_{\rm xc}(\mathbf{r}) = v_{\rm xc}(\mathbf{r})\mathsf{I} + \mu_B\boldsymbol{\sigma}\cdot\mathbf{B}_{\rm xc}(\mathbf{r}) \tag{28}$$

where $\mu_{\rm B} = \frac{e\hbar}{2mc}$ is the Bohr magneton.

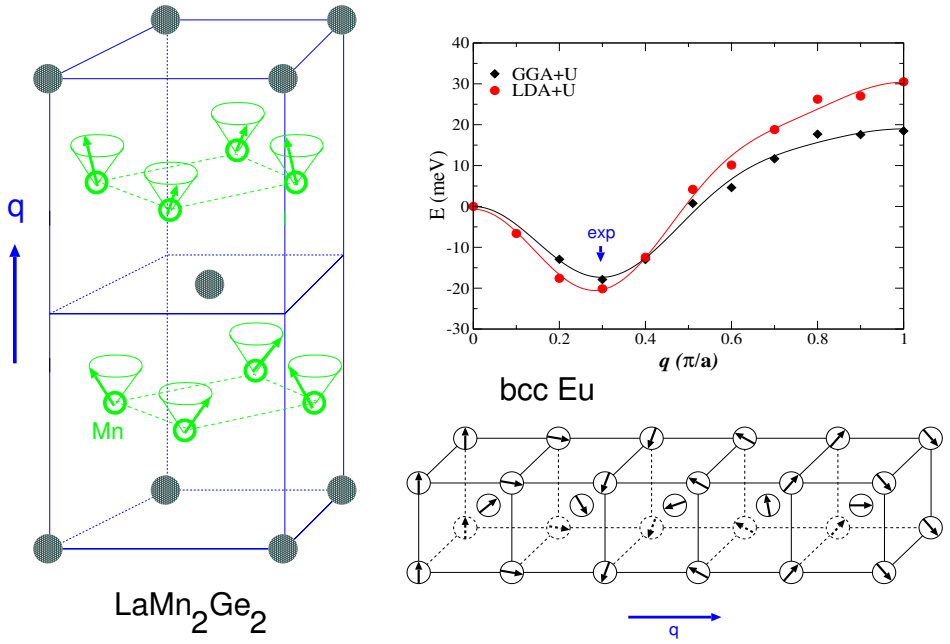


Fig. 12: *Examples of non-collinear ground-states: in $LaMn_2Ge_2$ the spins on the Mn sublattice can be described by a conical spin-spiral, i.e. the magnetization precesses on a cone with a semicone-angle of* $58°$ *from layer to layer in* z*-direction. The turning angle per layer is* $128°$*. The magnetic structure is in good agreement with first-principles calculations [63]. Bulk bcc europium has a flat spiral in* $[001]$ *direction as ground state (right, bottom), the length of the* $\mathbf{q}$*-vector, describing this precession, is correctly reproduced by DFT [64] (right, top).*

Within this formalism, general non-collinear structures can be described in the framework on density functional theory. Two recent examples, $LaMn_2Ge_2$ and bcc-Eu, are shown in Fig. 12.

Numerically, Eq. (25) is solved by expanding $\phi_i(\mathbf{r})$ in a linear combination of suitable basis-functions $\chi_j(\mathbf{r})$. Then Eq. (25) transforms into a eigenvalue problem and the eigenvectors, that have to be determined, give the linear combination coefficients, c_{ij}, of the expansion $\phi_i(\mathbf{r}) = \sum_j c_{ij}\chi_j(\mathbf{r})$. Such an eigenvalue problem is a standard problem of linear algebra and the computational effort scales in the most general case with the third power of the number of basis-functions. Compared to the non-magnetic problem, Eq. (23), this number is doubled in Eq. (25). Therefore, the computational effort for a general, non-collinear calculation is increased by a factor eight as compared to the non-magnetic calculation. An additional factor of at least 2 comes from the fact that the Hamiltonian is now hermitian and an other factor of about 4 can easily come in since the non collinear magnetism reduces the symmetry and a larger irreducible part of the Brillouine-zone has to be sampled.

3.2 The collinear case

Supposing that the potential matrices in Eq. (28) are diagonal (i.e. the magnetic and exchange fields point in z direction), Eq. (25) decouples into two equations of the type of Eq. (23):

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\rm Coul}(\mathbf{r}) + v(\mathbf{r}) + B_z(\mathbf{r}) + v_{\rm xc}^{(+)}(\mathbf{r})\right)\phi_i^{(+)}(\mathbf{r}) = \epsilon_i^{(+)}\phi_i^{(+)}(\mathbf{r}) \quad (29)$$
$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\rm Coul}(\mathbf{r}) + v(\mathbf{r}) - B_z(\mathbf{r}) + v_{\rm xc}^{(-)}(\mathbf{r})\right)\phi_i^{(-)}(\mathbf{r}) = \epsilon_i^{(-)}\phi_i^{(-)}(\mathbf{r})$$

where $v_{\rm Coul}$ denotes now the classical Coulomb potential and $v_{\rm xc}^{(+,-)}$ the exchange-correlation potential that arises from the functional derivative of the exchange-correlation energy with respect to the spin-up $(+)$ or spin-down $(-)$ part of the diagonal density matrix.
Since the two equations (29) can be solved independently, the computational effort for a collinear calculation seems to be just twice the effort for a non-magnetic calculation. However, most magnetic calculations are computationally considerably more demanding since the quantities in question (magnetic moments, energy differences between various magnetic configurations) require much higher accuracy than what is needed for nonmagnetic systems. To explore different magnetic orders in a system, unit cells much larger than the chemical unit cell are required, e.g. antiferromagnetic body-centered cubic (bcc) chromium requires a calculation with at least two atoms in the unit cell (as compared to one, in a nonmagnetic calculation).
Systems that can be described by Eq. (29) are all kinds of magnetic materials that assume a collinear magnetic order, e.g. ferromagnetic, antiferromagnetic or ferrimagnetic states. Like the density is a property that can – at least in principle – be obtained exactly in DFT, the spin density is a property that is well defined in spin-polarized DFT:

$$\mathbf{m}(\mathbf{r}) = -\mu_{\rm B}\sum_{\alpha,\beta}\psi_\alpha^+(\mathbf{r})\boldsymbol{\sigma}_{\alpha\beta}\psi_\beta(\mathbf{r}). \quad (30)$$

The integral spin moment, M, for a collinear system is then (in units of $\mu_{\rm B}$) simply

$$M_{\rm spin} = \int m(\mathbf{r})d\mathbf{r} = \mu_{\rm B}\int \left(n^{(+)}(\mathbf{r}) - n^{(-)}(\mathbf{r})\right)d\mathbf{r}. \quad (31)$$

How well this quantity corresponds to experimental values depends of course on the quality of the exchange-correlation potential that is used for an actual calculation. Some examples of results obtained in the local spin density approximation (LSDA) and generalized gradient approximation (GGA) of (spin)-moments of elemental ferromagnets are given in Table 1.
The success of DFT calculations for magnetic systems – as shown in Table 1 – has to be contrasted here with the case of Cr, where up to now no satisfactory agreement with experimental results was obtained: while LSDA calculations of antiferromagnetic Cr at the experimental lattice constant give a magnetic moment in nice agreement with experimental data ($0.5 - 0.6\mu_{\rm B}$), calculations at the lattice constant determined with LSDA (which is 3% too small) yield a nonmagnetic ground state. GGA calculations, on the other hand, give a reasonable lattice constant, but the magnetic moment is more than 60% too large [65]. Also attempts to include the experimentally observed incommensurate sinusoidal modulation of the antiferromagnetic structure of Cr could nor resolve these discrepancies so far.
At this point we should notice, that we relied on the assumption that the total energy is invariant with respect to a uniform rotation of the magnetization direction. This was implicitly assumed when we arbitrarily (or, better, for convenience) selected in Eq. (29) the z direction as global magnetization axis. Indeed, in absence of an external field (or in its presence, as long as it is oriented in z direction) this implies no loss in generality, if $v_{\rm xc}$ is isotropic in space. If we start

Table 1: *Magnetic moments (in μ_B per atom) of ferromagnetic elements in the bulk. The experimentally determined total magnetization, $M_{tot.}$, consists of spin- and orbital moment contributions. The LSDA results for Fe, Co and Ni are taken from Moruzzi et al. [66], the GGA values from Shallcross and coworkers [67] where also experimental values are quoted. The calculated Gd data is from Kurz et al. [68], the experimental one from White and coworkers [69].*

Property	source	Fe (bcc)	Co (fcc)	Ni (fcc)	Gd (hcp)
M_{spin}	LSDA	2.15	1.56	0.59	7.63
M_{spin}	GGA	2.22	1.62	0.62	7.65
M_{spin}	experiment	2.12	1.57	0.55	
$M_{tot.}$	experiment	2.22	1.71	0.61	7.63

from a Schrödinger-Pauli like theory, there is indeed no term that could couple the spin-space to the lattice. Only if a spin-orbit coupling term (from a Dirac type theory) or – in some cases – dipolar interaction is included, a preferential direction for the collinear magnetization exists. This will be discussed in the next section.

3.3 Orbital magnetism

The magnetization density we discussed in the last section, Eq. (30), is clearly a consequence of the imbalance of electrons with spin-up or spin-down and, therefore, the quantity defined in Eq. (31) is called spin-moment. From atomic physics we know, that the total magnetic moment is a sum of spin- and orbital contributions, $M_{tot.} = M_{spin} + M_{orb}$. The orbital moment results, in a classical picture, from the orbital motion of the electron around the nucleus. Compared to the situation in a free atom, where M_{orb} can be even larger than M_{spin}, in a solid this motion is of course restricted by the crystal field that quenches the orbital moment. In bulk samples small moments (typically $0.1 - 0.2\mu_B$) can be found (compare Table 1).
Density functional theory in the known LSDA or GGA formulations provides no term that could lead to the formation of an orbital moment. Current- and spin-density functional theory [70] would provide a natural starting point for the description of orbital magnetism, but so far the successes are limited. From relativistic quantum mechanics [71], a two-component approximation to the Dirac equation can be formulated, that has the form of Eq. (25). From the several terms appearing in this Pauli equation one term, the spin-orbit coupling term, provides a mechanism that leads to orbital polarization: The electron, traveling on a classical trajectory around the nucleus, experiences the electric field (from the screened nucleus) as a magnetic field. This field couples to the magnetic (spin) moment of the electron and, thus, leads to a preferential orientation of the orbital motion. Using the orbital angular moment operator $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, we can write the spin-orbit coupling term in the vicinity of a nucleus with a radial potential $v(r)$ as:

$$H_{so} = \frac{1}{r}\frac{dv(r)}{dr}(\boldsymbol{\sigma} \cdot \mathbf{L}). \tag{32}$$

Adding this term to Eq. (25) will destroy a decoupling of spin-up and -down equations like in Eq. (29). It also invalidates the aforementioned assumption that the total energy is not affected by a uniform rotation of the spin directions, since now spin-space and lattice are coupled by Eq. (32).

The orbital magnetization can be defined in analogy to Eq. (30), expressed in single particle wavefunctions ϕ_i:

$$\mathbf{m}(\mathbf{r}) = -\mu_{\rm B} \sum_i \langle \phi_i | \mathbf{r} \times \mathbf{p} | \phi_i \rangle. \tag{33}$$

At a certain atom ν, the orbital moment $M_\nu^{\rm orb}$ can then be obtained by an integration in a sphere centered around this atom:

$$M_\nu^{\rm orb} = -\mu_{\rm B} \sum_i \langle \phi_i | \mathbf{L} | \phi_i \rangle_\nu. \tag{34}$$

While this definition of the orbital moment poses no difficulties in periodic solids, we note here that the evaluation of the total orbital moment of a periodic crystal is more involved [72]. In most cases, however, the atomic orbital moments and also the magnetic anisotropy energies, obtained in density functional theory calculations, are too small as compared to experiment. Practical methods that can overcome this deficiency have been discussed in the literature [73].
Since the crystal field in a solid forces the orbital motion of the electron in a preferred plane, a total energy difference arises when the solid is magnetized in two different directions [74]. This difference, magnetocrystalline anisotropy energy (MAE), is small for bulk systems with high symmetry, e.g. cubic crystals like Fe or Ni. It is larger for crystals with a unique crystallographic axis, like hexagonal Co. But for lower dimensional systems, thin films or atomic wires, the magnetocrystalline anisotropy will essentially determine the magnetic properties, especially at finite temperatures [75].
The computational effort for calculations that include the spin-orbit coupling term, Eq. (32), can be reduced if this term is considered as a small perturbation to the non-relativistic Schrödinger-Pauli Hamiltonian. Then, the so-called magnetic force theorem [76] can be used to evaluate quantities like the MAE. But even these calculations require considerable computational resources, since the energy differences to be determined are very small and – compared to normal calculations – drastically increased numerical cutoffs can be necessary. Systems, where spin-orbit coupling is strong require a self-consistent treatment including Eq. (32) in the Hamiltonian. There, of course, the relativistic effects are stronger so that moderate numerical cutoffs can be used, but the computational complexity brought by the spin-orbit coupling term and the loss of symmetry that can be exploited leads to an increased computational effort.

4 The magnetic ground state

To determine the magnetic ground state it is possible to follow several directions: like in molecular-dynamics calculations, spin-dynamics allows to study the magnetic degrees of freedom exploring the ground state configuration. Another possibility is to determine the magnetic interactions between the atoms by a DFT calculation which are then mapped onto a model (in the simplest case a classical Heisenberg model). This model is then solved, either analytically or numerically. In both cases we introduce a discretization of the (vector) magnetization density: In spin-dynamics, the evolution of discrete spins, i.e. vectors attached to certain (atomic) positions is monitored. Also mapping the *ab initio* results to a model Hamiltonian which contains interactions between spins requires that it is possible to assign a definite spin to an atom, so that it should be possible to write in the vicinity of an atom ν, e.g. within some sphere centered at the nucleus, the magnetization density, $\mathbf{m}(\mathbf{r})$, as

$$\mathbf{m}(\mathbf{r}) = M_\nu \hat{\mathbf{e}}_\nu \tag{35}$$

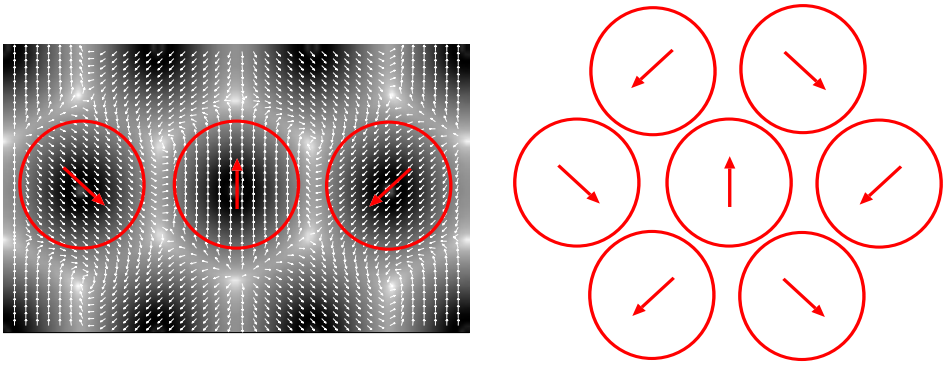


Fig. 13: *Left: ground state magnetization density of a hexagonal Cr monolayer with the Cu(111) in-plane lattice constant; the absolute value of the magnetization is shown in greyscale, the local directions are marked by small arrows. The average magnetization direction around an atom is indicated as red arrows. Right: schematic picture of the magnetic structure (Néel state) of the hexagonal Cr monolayer.*

where M_ν is the magnetization and $\hat{\mathrm{e}}_\nu$ is the magnetization direction. Vector-spin DFT calculations allow to estimate whether Eq. (35) is a good approximation or not (cf. Fig. 13). If the magnetization density in the vicinity of some atom ν is expressible by Eq. (35), then the total energy of a magnetic system as a function of its magnetic structure can be described as a functional $E[\{\hat{\mathrm{e}}_\nu\}]$ of the directions of the magnetic moments at the atoms ν in the magnetic unit cell. In this context collinear states ($\hat{\mathrm{e}}_\nu$ is identical for all atoms) are special solutions where $E[\{\hat{\mathrm{e}}_\nu\}]$ has a local or global maximum or minimum. Therefore, they constitute an important class of magnetic configurations that are often realized in magnetic materials. Unlike in non-spinpolarized DFT it is, however, in practical calculations not guaranteed that the obtained solution, $\mathsf{n}(\mathbf{r})$, is really the ground state and often several metastable solutions can be obtained.

4.1 Ab initio spin-dynamics, magnetic torque

If one is interested in the magnetic ground state of a system of given chemical composition and atomic positions, the final goal is to minimize the functional $E[\{\hat{\mathrm{e}}_\nu\}]$. The dimensionality of this problem will of course depend on the size of the chosen unit cell (some multiple of the chemical unit cell) and this minimization will involve the tricky task to determine the absolute minimum on a high-dimensional total energy surface. In analogy to molecular dynamics, i.e. the problem of minimizing the energy as a function of the atomic positions, we introduce here a spin dynamics, where the magnetic orientations, $\hat{\mathrm{e}}_\nu$, take the role of the variables.
Any vector-spin DFT calculation has to start with a reasonably chosen spin configuration in a prescribed unit cell. On a simple level, one can "relax" the directions of the magnetization at the atoms like a relaxation of the atomic structure (e.g. at a surface) is done. The magnetization directions, $\hat{\mathrm{e}}_\nu$, will then generally change to minimize the total energy (cf. Fig. 14). The final magnetic state, that will be reached, will in general depend on the starting point of the calculation and a more elaborate technique will be needed to avoid being trapped in some local minimum of $E[\{\hat{\mathrm{e}}_\nu\}]$.
To this end we have to develop an equation of motion for the magnetization of an atom. To keep things simple, we will focus on the case, where the magnetization stays collinear within the

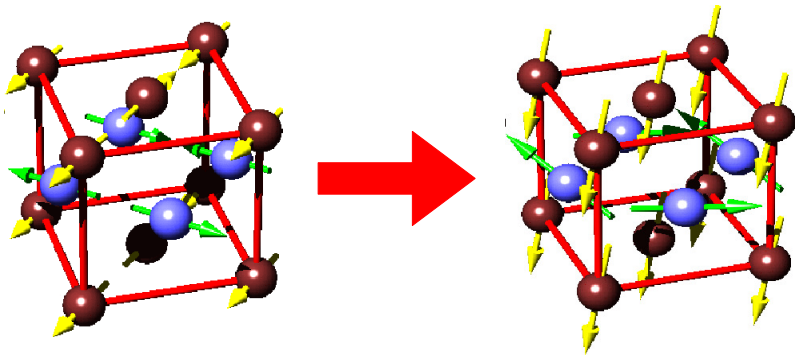


Fig. 14: *Determination of the magnetic ground state of ordered FeMn: the magnetic structure of the disordered alloy is a* $2\mathbf{q}$*-state (left). In an ordered alloy a more complex magnetic arrangement is obtained (right) by "relaxation" of the local spin directions.*

vicinity of the atom. Let us start from the Hamiltonian of Eq. (25) and assume that the external potential matrix, $\mathsf{v}(\mathbf{r})$, has been chosen to be diagonal and the exchange-correlation potential is separated into diagonal and off-diagonal parts. Following Antropov *et al.* [77, 78] we set up a time-dependent analogon of Eq. (25):

$$i\frac{d\Phi}{dt} = \left[H_d - \boldsymbol{\sigma}\cdot\mathbf{B}(\mathbf{r},t)\right]\Phi \quad \text{where} \quad \Phi = \begin{pmatrix} \phi^{(+)} \\ \phi^{(-)} \end{pmatrix}, \tag{36}$$

and H_d is the Hamiltonian that contains now only diagonal parts.
We will now separate the evolution of the magnetization into fast (value of the magnetization) and slow (direction of the magnetization) degrees of freedom. The former part will be described quantum-mechanically, while the latter is treated on a semiclassical level. At a given time, t, the time-independent version of Eq. (36) can be solved for a given magnetization characterized by $\{\hat{\mathbf{e}}_\nu\}$. Now we have to determine an equation of motion for the magnetization $\mathbf{m}(\mathbf{r},t)$.
This equation of motion can be obtained by multiplying Eq. (36) from the left with $\Phi^*\boldsymbol{\sigma}$ and adding the complex conjugate equation. Comparing to the time derivative of Eq. (30) and using the relation $\boldsymbol{\sigma}(\boldsymbol{\sigma}\cdot\mathbf{B}) = \mathbf{B} - i\boldsymbol{\sigma}\times\mathbf{B}$ we get

$$\frac{d\mathbf{m}(\mathbf{r},t)}{dt} = 2\mathbf{m}\times\mathbf{B} + \frac{i}{2}\nabla(\Phi^*\boldsymbol{\sigma}\cdot\nabla\Phi - c.c.). \tag{37}$$

The second term on the right side is complicated and describes longitudinal changes of the magnetization, which we will not consider on this level. Omitting this term, Eq. (37) describes the precession of the magnetization direction at an atom under the influence of the magnetic field generated by the atom itself and other atoms of the crystal.
Returning once more to Eq. (35), we can simplify Eq. (37) and write for the evolution of the magnetization direction in atom ν

$$\frac{d\hat{\mathbf{e}}_\nu}{dt} = -\frac{2}{\mu_\mathrm{B}}\hat{\mathbf{e}}_\nu\times\mathbf{I}_\nu \tag{38}$$

where $\mathbf{I}_\nu = \mu_\mathrm{B}\mathbf{B}$. If we explicitly also want to take into account the effect of other fields acting onto a magnetization direction, e.g. stemming from the spin-orbit interaction (magnetic anisotropy) or dipole-dipole interaction, these fields can be added to Eq. (38) into $\mathbf{I} = \mathbf{I}_\nu + \mathbf{I}_\mathrm{SO} +$

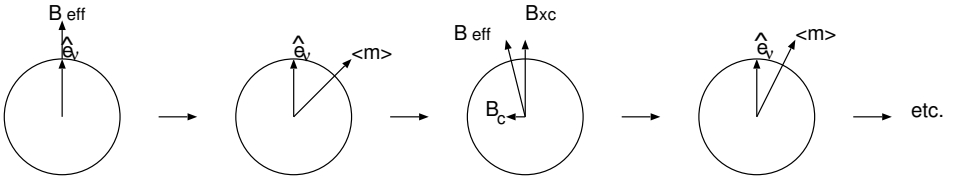


Fig. 15: *Determination of the constraint field: Initially, the effective* $\mathbf{B}$*-field is parallel to the prescribed direction* $\hat{\mathbf{e}}^{\nu}$ *(left). The resulting magnetization,* $\langle \mathbf{m} \rangle$*, generally is not parallel to this direction. Therefore, a constraint field* $\mathbf{B}_c$ *is introduced, that points in opposite direction to the component of the magnetization that is perpendicular to* $\hat{\mathbf{e}}^{\nu}$*. Using this* $\mathbf{B}_{\text{eff}}$*, the direction of the magnetization is then adjusted towards* $\hat{\mathbf{e}}^{\nu}$ *(right).*

$\mathbf{I}_{\mathrm{d-d}}$. More general expressions of Eq. (38), suitable for spin-dynamics with finite temperatures included, can be found in reference [78].
The next question, that has to be answered, is how to determine the fields $\mathbf{I}_{\nu}$, i.e. given a certain set of magnetization directions $\{\hat{\mathbf{e}}_{\nu}\}$ what is the torque on a selected magnetic moment [79]. This problem can be solved in constrained vector-spin density functional theory, as introduced in the next section.

4.2 Constrained density functional theory

In general, an arbitrary magnetic configuration given by a set of local (atomic) magnetization directions $\{\hat{\mathbf{e}}_{\nu}\}$ is not an extremum or a stationary solution of the total energy functional $E\left[\mathsf{n}(\mathbf{r})\right]$. Exceptions are high symmetry states, like collinear magnetic states, a certain class of spin-spiral states (see Sec. 4.4) and particular linear superpositions of several spin-spiral states. The constrained density functional theory developed by Dederichs *et al.* [80] provides the necessary generalization to deal with arbitrary magnetic configurations, i.e. configurations where the orientations of the local moments are constrained to non-equilibrium directions. We define a generalized energy functional $\tilde{E}\left[\mathsf{n}(\mathbf{r})|\{\hat{\mathbf{e}}_{\nu}\}\right]$, where we ensure that the average magnetization in an atom, $\langle \mathbf{m} \rangle_{\nu}$, points in the direction $\hat{\mathbf{e}}_{\nu}$. This condition, $\hat{\mathbf{e}}_{\nu} \times \langle \mathbf{m} \rangle_{\nu} = 0$, is introduced by a Lagrange multiplier, $\boldsymbol{\lambda}$, so that [7]

$$\begin{aligned} \tilde{E}\left[\mathsf{n}(\mathbf{r})|\{\hat{\mathbf{e}}_{\nu}\}\right] &= E\left[\mathsf{n}(\mathbf{r})\right] + \sum_{\nu} \boldsymbol{\lambda}^{\nu} \cdot \left(\hat{\mathbf{e}}_{\nu} \times \langle \mathbf{m} \rangle_{\nu}\right) \\ &= E\left[\mathsf{n}(\mathbf{r})\right] + \mu_B \sum_{\nu} \mathbf{B}_c^{\nu} \cdot \langle \mathbf{m} \rangle_{\nu}. \end{aligned} \tag{39}$$

Here, we recast the Lagrange multiplier in the form of a magnetic field, $\mathbf{B}_c^{\nu}$, which is the constraining field in atom ν that keeps the local (integrated) magnetic moment, i.e. the magnetization density averaged over the sphere where Eq. (35) holds,

$$\langle \mathbf{m}(\mathbf{r}) \rangle_{\nu} = \mathbf{M}^{\nu} = \int_{MT^{\nu}} \mathbf{m}(\mathbf{r})\, d^3r, \tag{40}$$

parallel to the prescribed direction $\hat{\mathbf{e}}_{\nu}$.
In an actual constrained local moment (CLM) calculation $\mathsf{n}(\mathbf{r})$ and $\mathbf{B}_c^{\nu}$ have to be determined self-consistently. The density matrix is calculated in the usual self-consistency cycle. At the

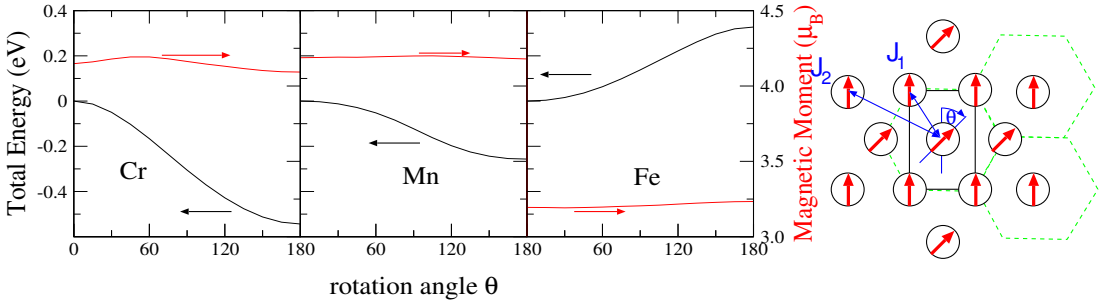


Fig. 16: *Total energy and magnetic moment of hexagonal monolayers of Cr, Mn, and Fe as a function of the angle of the magnetization in a two-atomic unit cell (right). $\theta = 0°$ corresponds to a ferromagnetic state, $\theta = 180°$ is a row-wise antiferromagnetic state. As lattice constant we chose the parameters of the Ag(111) surface. In the schematic picture of the hexagonal monolayer (right) the coupling to nearest neighbors (J_1) and next- nearest neighbors (J_2) is indicated.*

same time, the local constraint fields $\mathbf{B}_c^\nu$ have to be adjusted, until the constraint conditions, $\hat{\mathbf{e}}_\nu \times \langle \mathbf{m} \rangle_\nu = 0$, are fulfilled (cf. Fig. 15). At the end of such a calculation we obtain the self-consistent densities and a set of local constraint $\mathbf{B}$-fields. The total energy of the system is given by the constrained energy functional, Eq. (39).
According to the the Hellmann-Feynman theorem we find that the change of the energy due to a change in magnetization direction, $d\hat{\mathbf{e}}_\nu$, is given by $dE = -\mu_B \mathbf{M}^\nu \cdot (\mathbf{B}_c^\nu \times d\hat{\mathbf{e}}_\nu)$. Therefore, the constraint field can be interpreted as a torque acting on the magnetic moment, in the spirit of the spin dynamics, formulated in the previous section. Thus, we have set up a formalism that allows us to find – at least in principle – the magnetic ground state of a system by spin-dynamics [81]. But CLM calculations can also be used in a different way: In the next section we will describe how they can be used to determine the exchange interactions in a system and utilize these results in models, like the classical Heisenberg model, to obtain information about the ground state, but also about excited states of a magnetic system.

4.3 Mapping onto realistic model Hamiltonians: canted moments

The classical Heisenberg model (2) can be used as a phenomenological starting point for the investigation of the magnetic interaction in a crystal. Although the Heisenberg model was originally introduced for magnetic insulators with localized moments [82], we can also apply Eq. (2) to metallic systems, as shown in Fig. 16. In these hexagonal unsupported monolayers the behavior of the total energy as a function of the relative angle between the atoms can be described as cosine-like function, the exchange coupling constant being negative for Cr and Mn (preferring antiferromagnetic coupling) and positive for Fe (leading to a ferromagnetic ground state). The total energy has been calculated by a constrained DFT calculation as described above. We further see, that the magnetic moment does not change significantly as the spins are rotated, an important requirement for the application of the Heisenberg model.
From the right part of Fig. 16 we can see, that rotating the local magnetic moment direction of one atom in the two-atom unit cell of the hexagonal lattice will change the relative orientation of that atom to four nearest neighbors, but does not affect two of the nearest neighbor (NN) atoms. Likewise, only four of the six second-NN atoms will change the relative orientation to

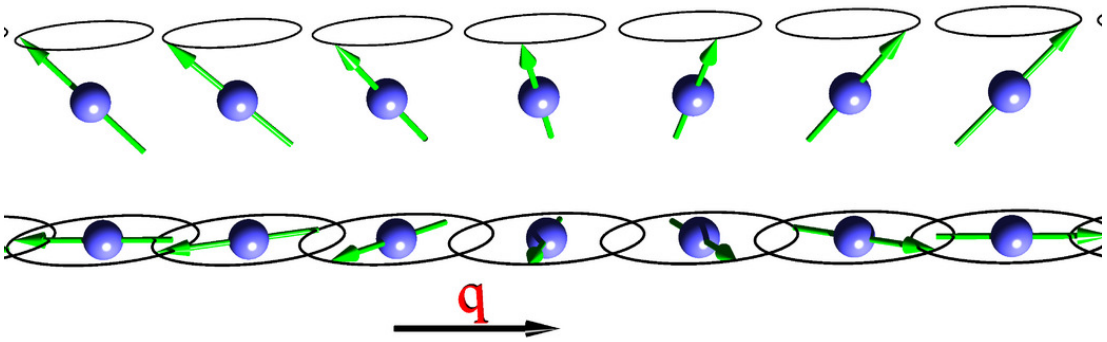


Fig. 17: *Spiral spin-density wave (SSDW) or spin spiral propagating along the with a wavevector* $\mathbf{q}$*. The upper spiral represents a coned-spiral with a cone angle of* $45°$*, the bottom spiral represents a flat spiral, with a cone angle of* $90°$*. Both spirals exhibit a counterclockwise rotational sense expressed by a chirality vector parallel to the* z*-axis.*

the original atom. This leads to an expression for the total energy in the classical Heisenberg model up to second-NN:

$$E = -S^2(J_1 + J_2)(2 + 4\cos\theta) \tag{41}$$

if $\mathbf{S}$ is now the total spin moment treated as a classical vector. This means, from a constrained local moment calculation we can at least estimate the size of $(J_1 + J_2)$. It is not difficult to find other unit cells and rotations that allow the determination of other linear combinations of J_1 and J_2, thereby separating the individual exchange coupling constants [53].
Of course, the energies obtained from the CLM calculation contain contributions of all J_n and also from interactions that are not described by the Heisenberg model. Examples, like the biquadratic interaction or the 4-spin interaction result from hopping processes between four sites, inclusion of spin-orbit interaction gives rise to a third-order process, the so-called Dzyaloshinsky-Moriya interaction [83]. All these different interaction terms can be extracted from a set of suitable *ab initio* calculations (possibly including spin-orbit interaction) and can be used to determine the magnetic ground state within the chosen model.

4.4 Mapping onto realistic model Hamiltonians: spin-spirals

Concluding from the previous section, one way to obtain model parameters like the elements of the above discussed exchange parameters J_{ij} from first principles is to fit the energy expression obtained from the model ansatz to the total energies obtained from electronic structure calculations for different magnetic states. In a periodic crystal it is convenient to replace the quantities in Eq. (2) by their Fourier-transformed equivalents:

$$\mathbf{S}(\mathbf{q}) = \frac{1}{N}\sum_n \mathbf{S}_n e^{-i\mathbf{q}\mathbf{R}_n} \quad \text{and} \quad J(\mathbf{q}) = \sum_n J_{0n} e^{-i\mathbf{q}\mathbf{R}_n}. \tag{42}$$

Exploiting the translational invariance of the lattice, we can then rewrite Eq. (2) as

$$H = -N\sum_{\mathbf{q}} J(\mathbf{q})\mathbf{S}(\mathbf{q})\cdot\mathbf{S}(-\mathbf{q}) \tag{43}$$

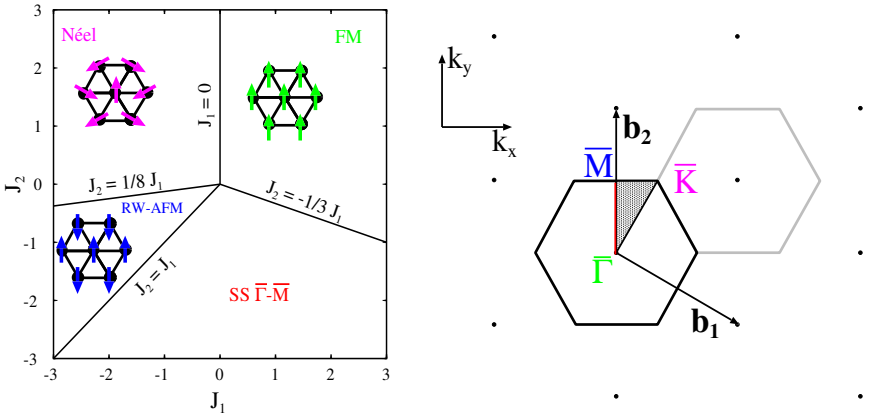


Fig. 18: *Zero temperature phase-diagram of the classical Heisenberg model for the hexagonal lattice in next-nearest neighbor approximation of the classical Heisenberg model: two collinear, a ferromagnetic (FM) and a row-wise antiferromagnetic (RW-AFM) solution can be obtained, and two non-collinear solutions, the Néel-state and a SSDW with* $\mathbf{q}$*-vectors along the line* $\overline{\Gamma}-\overline{\mathrm{M}}$ *of the Brillouin zone (right). The FM, RW-AFM and Néel-state correspond to SSDWs with* $\mathbf{q}$*-vectors on the high-symmetry points* $\overline{\Gamma}$*,* $\overline{\mathrm{M}}$ *and* $\overline{\mathrm{K}}$*, respectively.*

where we have to ensure that the length of all spins $\mathbf{S}_n^2 = S^2$ is conserved on all sites n. This condition is fulfilled by solutions of the type [83]

$$\mathbf{S}_n = S\left(\hat{\mathbf{e}}_x \cos(\mathbf{q}\cdot\mathbf{R}_n) + \hat{\mathbf{e}}_y \sin(\mathbf{q}\cdot\mathbf{R}_n)\right) \tag{44}$$

where the unit vectors $\hat{\mathbf{e}}_x$ and $\hat{\mathbf{e}}_y$ just have to be perpendicular to each other, otherwise their directions are arbitrary. Eq. (44) describes a spiral spin-density wave (SSDW) as shown in the lower half of Fig. 17. A more generals form of SSDWs can be obtained, when the magnetization precesses on a cone with an opening angle ϑ:

$$\mathbf{S}_n = S\left(\hat{\mathbf{e}}_x \cos(\mathbf{q}\cdot\mathbf{R}_n)\sin\vartheta + \hat{\mathbf{e}}_y \sin(\mathbf{q}\cdot\mathbf{R}_n)\sin\vartheta + \hat{\mathbf{e}}_z\cos\vartheta\right) \tag{45}$$

as shown in the upper half of Fig. 17.
These SSDWs are general solutions of the classical Heisenberg model for a periodic lattice. From Eq. (43) one can conclude that the SSDW with the lowest total energy will be the one with the propagation vector $\mathbf{Q}$ that maximizes $J(\mathbf{q})$. These will be preferably the high-symmetry points in the Brillouin-zone of $\mathbf{q}$-vectors, and then high-symmetry lines. For example, if $\mathbf{Q}=0$ maximizes $J(\mathbf{q})$, the solution corresponds to the ferromagnetic state, if $\mathbf{Q} = \hat{\mathbf{e}}_z\frac{\pi}{a_z}$ and a_z is the lattice constant in z-direction, then the structure is layered antiferromagnetic in z-direction. Some other examples – for a hexagonal monolayer – are illustrated in Fig. 18. $T=0$ many compounds and elemental metals show SSDW ground states. Some examples were shown in Fig. 12.

4.5 Spin-spirals and the generalized Bloch theorem

A very elegant treatment of spin-spirals by first-principle calculations is possible when the generalized Bloch theorem [84, 85, 86] is applied. However, this theorem can only be proved, when spin-orbit coupling is neglected. For this reason the spin-rotation axis will always be considered

as parallel to the z-axis of the spin coordinate-frame. Thus, only the m_x and m_y components are rotated, while m_z does not change. Following Sandratskii [85] we can define a generalized translation, $\mathcal{T}_n$, that combines a lattice translation, $\mathbf{R}_n$, and a spin rotation $\mathbf{U}$ that commutes with the Hamiltonian $\mathcal{H}$. Applying such a generalized translation to $\mathcal{H}\Phi$ yields

$$\begin{aligned} \mathcal{T}_n\mathcal{H}(\mathbf{r})\Phi(\mathbf{r}) &= \mathbf{U}(-\mathbf{q}\mathbf{R}_n)\mathcal{H}(\mathbf{r}+\mathbf{R}_n)\mathbf{U}^\dagger(-\mathbf{q}\mathbf{R}_n)\mathbf{U}(-\mathbf{q}\mathbf{R}_n)\Phi(\mathbf{r}+\mathbf{R}_n) \\ &= H(\mathbf{r})\mathbf{U}(-\mathbf{q}\mathbf{R}_n)\Phi(\mathbf{r}+\mathbf{R}_n) \end{aligned} \tag{46}$$

where $\mathbf{U}(\mathbf{q}\mathbf{R}_n)$ is the spin $1/2$ rotation matrix

$$\mathbf{U}(\mathbf{q}\mathbf{R}_n) = \begin{pmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix}, \quad \varphi = \mathbf{q}\cdot\mathbf{R}_n. \tag{47}$$

In analogy with the proof of Bloch's theorem [87] it follows that the eigenstates can be chosen such that

$$\mathcal{T}_n\Phi(\mathbf{k},\mathbf{r}) = \mathbf{U}(-\mathbf{q}\mathbf{R}_n)\Phi(\mathbf{k},\mathbf{r}+\mathbf{R}_n) = e^{i\mathbf{k}\cdot\mathbf{R}_n}\Phi(\mathbf{k},\mathbf{r}). \tag{48}$$

Since these eigenstates are labeled with the same Bloch vector $\mathbf{k}$ as the eigenstates of the translation operator without the spin rotation, the lattice periodic part of these states follows the chemical lattice, $\mathbf{R}_n$, i.e. we can calculate the spin spiral state in the chemical unit cell and $\mathbf{k}$ denotes a $\mathbf{k}$-vector of the reciprocal chemical lattice. The resulting eigenstates of the Hamiltonian $\mathcal{H}$ have the form

$$\psi_{\mathbf{k},\nu}(\mathbf{r}) = \begin{pmatrix} \psi^{(\uparrow)}_{\mathbf{k},\nu}(\mathbf{r}) \\ \psi^{(\downarrow)}_{\mathbf{k},\nu}(\mathbf{r}) \end{pmatrix} = \exp(i\,\mathbf{k}\cdot\mathbf{r}) \begin{pmatrix} \exp(-i\,\frac{1}{2}\,\mathbf{q}\cdot\mathbf{r})\; u^{(\uparrow)}_{\mathbf{k},\nu}(\mathbf{r}) \\ \exp(+i\,\frac{1}{2}\,\mathbf{q}\cdot\mathbf{r})\; u^{(\downarrow)}_{\mathbf{k},\nu}(\mathbf{r}) \end{pmatrix} \tag{49}$$

with the $\mathbf{q}$-vector of length $|q|$ pointing along the propagation direction of the spiral and the functions $u^{(\uparrow)}_{\mathbf{k},\nu}(\mathbf{r})$, $u^{(\downarrow)}_{\mathbf{k},\nu}(\mathbf{r})$ possessing the period of the chemical lattice.

In a reciprocal-space method, i.e. when all quantities like potential or wavefunctions are expressed as Fourier-transforms, the computational effort scales with the size of the unit cell. Without the application of the generalized Bloch theorem the investigation of spin spiral states requires very large unit cells, and a description of SSDWs that are incommensurate with the lattice would be not possible.

Since the spin-spiral is the exact solution of the classical Heisenberg model at T$=0$, it is believed that they cover a large and important part of the phase space of possible spin states. Thus among all possible magnetic states, spin-spirals are the next relevant class of spin states besides the high-symmetry magnetic states, i.e. the ferromagnetic, antiferromagnetic, or ferrimagnetic configurations.

A further computational simplification can be reached, when the SSDW is considered just as a small perturbation to the parent (most often ferromagnetic) structure. This may be justified in the limit of small $\mathbf{q}$-vectors or small opening angles ϑ (cf. Eq. (45)). The limit of $\vartheta \to 0$ is particularly important in the study of finite temperature effects, since it describes elementary perturbations of the collinear ground state. In this limit again the magnetic force theorem [76] can be applied, thus reducing the computational efforts significantly [88].

In real-space methods the calculation of $J(\mathbf{q})$ is most conveniently done via the right of Eq. (42), i.e. the evaluation of J_{0n}. In this case the direction of the magnetization at a reference atom, 0, is perturbed and the response on the other atoms, n, calculated. Also in this case a kind of magnetic force theorem can be used [64].

4.6 Calculating the DM interaction with periodic spin spirals

In order to determine the DM-strength $\mathbf{D}$ from first principles we follow the same strategy as above and fit an appropriate energy expression obtained from the model ansatz to the total energies obtained from electronic structure calculations for different magnetic states. We start with the model ansatz (18) and assume a situation where the system is in a collinear magnetic state without DM-interaction, the DM-interaction is much smaller than the exchange ($D << J$) and thus the DM-interaction can only introduce long-period spirals. Under this assumption we can go from a model on a discrete model (18) to a micromagnetic model that is continuous in space coordinates. We assume further that the easy axis is out-of-plane $\mathbf{K} = (0, 0, K)$ and the the hard axis is oriented parallel to $\mathbf{D} = (0, D, 0)$, then the anisotropy term also favors a magnetization that is confined to the plane normal to $\mathbf{D}$. In such a system with $\mathbf{m} \perp \mathbf{D}$, we can describe the magnetization as a planar spiral with only one angle φ ($\mathbf{m} = \hat{\mathbf{e}}_x \cos\varphi + \hat{\mathbf{e}}_z \sin\varphi$). For homogeneous spin spirals (i.e. spirals with $\dot{\varphi}(x) = q = \text{const}$ or a constant canting angle $\arccos(\mathbf{S}_i \cdot \mathbf{S}_{i+1})$ between the magnetizations of two adjacent lattice sites) the magnetic structure has a period of $\lambda_{\text{hs}} = 2\,\pi/q$ and the energy density E/λ_{hs} of the micromagnetic model gets

$$\frac{q}{2\,\pi}\,E(q) = \frac{q}{2\,\pi}\int_0^{\frac{2\,\pi}{q}} \mathrm{d}x\,\left(\,A\,q^2 + D\,q + K\sin^2(q\,x)\,\right) = A\,q^2 + D\,q + \tfrac{1}{2}\,K\;. \tag{50}$$

Thus, we can obtain our model parameters A for the exchange and D from a quadratic and a linear fit to the dispersion curve $q\,E(q)$, with latter obtained from first-principles calculations. In order to determine the anisotropy constant K, we can perform independent calculations of collinear configurations with $\varphi = 0$ and $\varphi = \pi$.

Since the micromagnetic model is valid in the limit of slow spatial rotations, relation (50) holds only for homogeneous spin spirals with large period lengths. This presents a formidable problem, since the size of the unit cell that one can treat is limited by the computing facilities. The generalized Bloch theorem [84, 85, 86] cannot be applied here directly due to the presence of the spin-orbit interaction. We employ a perturbative scheme in order to deal with these large magnetic superstructures: We calculate the rotations self-consistently on the basis of the Hamiltonian $\mathcal{H}_0$, but treat the spin-orbit coupling as a perturbation. Thereby, we make use of the local force theorem [89, 90]. If we neglect spin-orbit coupling, the orientation of the magnetic moments with respect to the crystal lattice is irrelevant and we can calculate the electronic structure of a homogeneous spin spiral within the chemical unit cell by applying a generalized Bloch theorem [84, 85, 86].

In a next step, we apply the spin-orbit coupling operator $\mathcal{H}_{\text{so}}$ in second variation, i.e. we expand the eigenfunctions of $\mathcal{H}_0 + \mathcal{H}_{\text{so}}$ in eigenfunctions of $\mathcal{H}_0$ and construct the Hamiltonian matrix with the matrix elements $\langle\psi_{\mathbf{k}',\nu'}|\mathcal{H}_0 + \mathcal{H}_{\text{so}}|\psi_{\mathbf{k},\nu}\rangle$. Since we can neglect all states of high energy, this procedure reduces the size of the Hamiltonian matrix drastically. The spin-orbit coupling is well described by

$$\mathcal{H}_{\text{so}} = \sum_\alpha \frac{1}{r_\alpha}\frac{\mathrm{d}V_\alpha(r_\alpha)}{\mathrm{d}r_\alpha}\,\boldsymbol{\sigma}\cdot\hat{\mathbf{L}}_\alpha = \begin{pmatrix} \mathcal{H}_{\text{so}}^{(\uparrow,\uparrow)} & \mathcal{H}_{\text{so}}^{(\uparrow,\downarrow)} \\ \mathcal{H}_{\text{so}}^{(\downarrow,\uparrow)} & \mathcal{H}_{\text{so}}^{(\downarrow,\downarrow)} \end{pmatrix}\,, \tag{51}$$

where the index α denotes the atoms. Since only the spin-independent part of the potential enters $\mathcal{H}_{\text{so}}$, its real-space representation possesses the period of the chemical lattice. This allows

us to write the matrix elements in the form

$$\begin{aligned}
\langle\psi^{(\uparrow)}_{\mathbf{k}',\nu'}|\,\mathcal{H}^{(\uparrow,\uparrow)}_{\mathrm{so}}\,|\psi^{(\uparrow)}_{\mathbf{k},\nu}\rangle &= \int\mathrm{d}\mathbf{r}\ \mathrm{e}^{i\,(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}\quad u^{(\uparrow,\uparrow)}(\mathbf{r})\ ,\\
\langle\psi^{(\uparrow)}_{\mathbf{k}',\nu'}|\,\mathcal{H}^{(\uparrow,\downarrow)}_{\mathrm{so}}\,|\psi^{(\downarrow)}_{\mathbf{k},\nu}\rangle &= \int\mathrm{d}\mathbf{r}\ \mathrm{e}^{i\,(\mathbf{k}-\mathbf{k}'+\mathbf{q})\cdot\mathbf{r}}\ u^{(\uparrow,\downarrow)}(\mathbf{r})\ ,\\
\langle\psi^{(\downarrow)}_{\mathbf{k}',\nu'}|\,\mathcal{H}^{(\downarrow,\uparrow)}_{\mathrm{so}}\,|\psi^{(\uparrow)}_{\mathbf{k},\nu}\rangle &= \int\mathrm{d}\mathbf{r}\ \mathrm{e}^{i\,(\mathbf{k}-\mathbf{k}'-\mathbf{q})\cdot\mathbf{r}}\ u^{(\downarrow,\uparrow)}(\mathbf{r})\ ,\\
\langle\psi^{(\downarrow)}_{\mathbf{k}',\nu'}|\,\mathcal{H}^{(\downarrow,\downarrow)}_{\mathrm{so}}\,|\psi^{(\downarrow)}_{\mathbf{k},\nu}\rangle &= \int\mathrm{d}\mathbf{r}\ \mathrm{e}^{i\,(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}\quad u^{(\downarrow,\downarrow)}(\mathbf{r})
\end{aligned}\tag{52}$$

with lattice-periodic functions $u^{(\uparrow,\uparrow)}$, $u^{(\uparrow,\downarrow)}$, $u^{(\downarrow,\uparrow)}$, $u^{(\downarrow,\downarrow)}$. Obviously, these matrix elements are non-zero if and only if the exponents are zero. If we choose a $\mathbf{q}$-vector that is commensurate with the reciprocal lattice (i.e. a spiral that is commensurate within a certain supercell) and a $\mathbf{k}$-grid that is commensurate to the $\mathbf{q}$-vector, then we obtain a block-diagonal and sparse Hamiltonian matrix. Each block contains the matrix elements from all $\mathbf{k}$-vectors that are connected by $\mathbf{q}$ (cf. Fig. 19). Note, that $\mathbf{q}$ is a reciprocal lattice vector of the large magnetic unit cell in which the spin spiral is commensurate.
For large systems, the Hamiltonian matrix that is shown in Fig. 19 is too large for straightforward diagonalization. But since we are applying the local force theorem, we only need to know the sum of occupied eigenvalues. This allows to use a perturbative technique that requires exact diagonalization only in a subspace close to the Fermi energy and exploits the sparseness [91].

5 Beyond the ground state

5.1 Low temperatures: magnons and spin waves

SSDWs are sometimes also called frozen magnons, since a spin-spiral looks like a "snap shot" of a single magnon at a fixed time. Spin-spiral calculations can therefore be used to simulate

$$H_{\{\mathbf{k}\}} = \Big(\ \delta_{\mathbf{k},\mathbf{k}'}\delta_{\nu,\nu'}\,\epsilon_{\mathbf{k},\nu} + \langle\psi_{\mathbf{k},\nu'}|\mathcal{H}_{\mathrm{so}}|\psi_{\mathbf{k},\nu}\rangle\ \Big)_{\substack{\mathbf{k},\mathbf{k}'\\ \nu,\nu'}} =$$

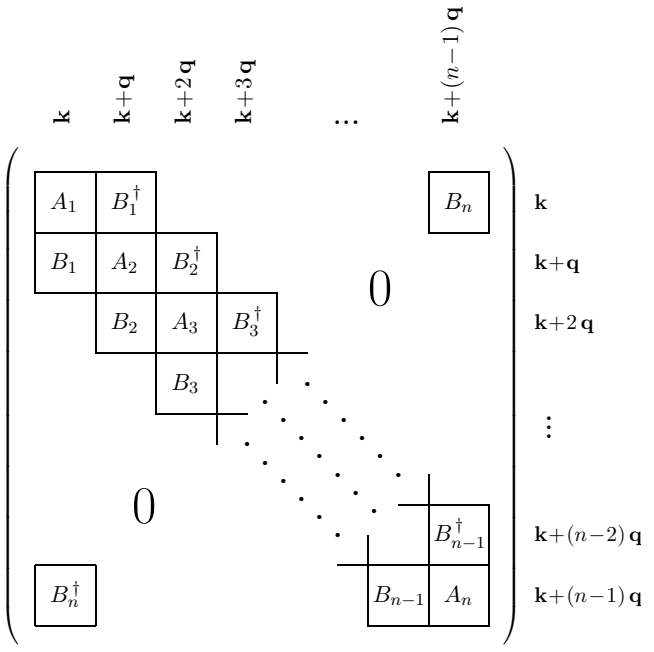


Fig. 19: *Block of the Hamiltonian matrix for spin-spiral basis functions.*

the effect of temperature on a magnetic system in the adiabatic approximation, in particular at very low temperatures, when magnons with long wavelength dominate.
At low, but finite temperatures, collective spin-wave excitations or magnons are excited in the ferromagnetic crystal. These magnons can again be characterized by their wave-vector $\mathbf{q}$. In the long wavelength limit, i.e. around $\mathbf{q} = 0$ the spin-wave dispersion behaves almost quadratically and can be described as Dq^2. The spin stiffness, D, characterizes the magnetic properties of a ferromagnet at low temperatures and can also be calculated from the exchange coupling constants:

$$D = \frac{2}{3M} \sum_n J_{0n} R_{0n}^2. \tag{53}$$

Here, M is the magnetic moment of the ferromagnetic state. Some results of *ab initio* calculations are given in Table 2. For Fe and Co agreement with experimental data is reasonable, but

Table 2: *Calculated and experimental spin-wave stiffness (D) for Fe, Co and Ni. The theoretical data was obtained in different approximations as described by Rosengaard and Johansson [92] [th.(1)], Shallcross and coworkers [67] [th.(2)] and Pajda et al. [93] [th.(3)], experimental data was taken as cited in these references.*

	D (meV Å^2)			
	th.(1)	th.(2)	th.(3)	exp.
Fe (bcc)	247	322, 313	250	280, 314, 330
Co (fcc)	502	480, 520	663	510, 580
Ni (fcc)	739	541, 1796	756	422, 550, 555

for Ni most methods fail to reproduce the experimental spin stiffness.

5.2 High temperatures: $T_{\rm C}$ and $T_{\rm N}$

Let us now see, how higher temperatures will influence the magnetic order in a ferromagnetic solid. Staying within the Heisenberg model, we will assume that the magnitude of the magnetic moments at the atoms will – in first approximation – not be changed, and discuss just their mutual orientation. At $T = 0$ the spin at a selected atom will be fixed in parallel direction to the spins at all other atoms by an effective field that will be proportional to $S \sum_n J_{0n} = S J_0$. At a finite temperature T, this field, that acts on the spin at site 0 is reduced due to the thermal fluctuation on the sites n. The thermal average of the projection of the spin at site n on the spin at site 0 is denoted as $\langle S(\mathbf{R}_n)\rangle$. In the "mean field approximation" (MFA), it is assumed that the effective field at finite temperatures that acts on spin 0 is:

$$B_{\rm eff} = \sum_n J_{0n} \langle S(\mathbf{R}_n)\rangle \tag{54}$$

In this model it is possible to calculate the temperature-dependence of the average magnetization of the solid and, specifically, the temperature where the average magnetization vanishes, the critical temperature. For a ferromagnet this temperature is called Curie temperature and in the MFA it is given by

$$T_{\rm C} = \frac{2S(S+1)}{3k_{\rm B}} J_0 \tag{55}$$

It has to be mentioned, that in most cases the MFA severely overestimates T_{C} (by about 20 to 50%, depending on the lattice). Nevertheless, it gives a simple estimate of the ordering temperature in systems, where the approximations of the Heisenberg model are reasonable. On the other hand, some properties, like the – material independent – critical exponents, are in any case not usefully reproduced by the MFA.
On a more sophisticated level, the "random phase approximation" (RPA) can give quite reliable results. In contrast to the MFA, where the thermal averaging was done over the sites n that determine B_{eff}, here the Hamiltonian is first transformed into reciprocal space (Eq. (43)), and then the averaging is done over one of the Fourier components:

$$H = -N \sum_{\mathbf{q}} J(\mathbf{q}) \mathbf{S}(\mathbf{q}) \cdot \langle \mathbf{S}(-\mathbf{q}) \rangle \tag{56}$$

If the term $S(S+1)$ is included in the exchange coupling constants (as it is usually done, when the J's are determined from first-principles calculations), then the Curie temperature in the MFA and RPA can be expressed as

$$k_{\mathrm{B}} T_{\mathrm{C}}^{\mathrm{MFA}} = \frac{2}{3} J_0 \qquad k_{\mathrm{B}} T_{\mathrm{C}}^{\mathrm{RPA}} = \frac{2}{3} \left(\sum_{\mathbf{q}} \frac{1}{J(\mathbf{q})} \right)^{-1} \tag{57}$$

From these expressions it is obvious, that calculating T_{C} in the RPA involves not more information that what is needed on a mean-field level, if the exchange coupling constants are calculated in reciprocal space by using the generalized Bloch theorem.
Also for antiferromagnets (or, generally spin-spiral states characterized by a vector $\mathbf{Q}$) expressions for the ordering temperature, the Néel temperature T_{N}, can be derived. In the MFA with $S(S+1)$ again included in J, this is given simply by

$$k_{\mathrm{B}} T_{\mathrm{N}}^{\mathrm{MFA}} = \frac{2}{3} J(\mathbf{Q}) \tag{58}$$

while a slightly more involved expression can be derived in the random phase approximation [64]. Comparison of these results with experimental values gave reasonable results, e.g. for bcc europium Néel temperatures of 147 K and 110 K were obtained in MFA and RPA, respectively [64]. These values have to be compared to the experimental T_{N} of 90.5 ± 0.5 K.
Although there exist several more methods to calculate critical temperatures from DFT results, we will outline here just one further possibility, which seems to be rather flexible and appropriate for many systems with different magnetic ground states: the Monte Carlo technique (MC) allows to study finite-temperature magnetic properties by implementation of a Heisenberg Hamiltonian (Eq. (2), possible with extensions like biquadratic terms or an uniaxial anisotropy (see below)), into a Metropolis algorithm [94]. Unit cells of different size are then studied so that finite-size effects can be eliminated. In these unit cells the evolution of the magnetic property in question (in our case the average magnetization) as a function of temperature can then be monitored. The method will be discussed in detail in Lecture C1 *Magnetic phase transitions: from density functional theory to Monte Carlo simulations.*
Results of *ab initio* calculations of the Curie temperature of Fe, Co and Ni are presented in Table 3. From this table one can easily see that, compared to RPA, the MFA typically overestimates T_{C} by 25–50%. For Fe and Co RPA gives quite good estimates of the Curie temperature, while for Ni T_{C} is underestimated in both approximations. MC simulations work better for Ni and Fe, but give a too low T_{C} for Co.

Table 3: *Calculated and experimental Curie temperature $T_{\rm C}$ for some ferromagnetic materials. MFA and RPA data for Fe, Co and Ni taken from Pajda et al. [93], MFA2 results and experimental values as quoted by Shallcross and coworkers [67], while the MC results were obtained by Rosengaard and Johansson [92]. Data for Gd can be found in the papers of Kurz et al. [68] and Turek and coworkers [95].*

	$T_{\rm C}$ (K)				
	MFA	MFA2	RPA	MC	exp.
Fe (bcc)	1414	$550, 1190$	950	1060	$1044 - 1045$
Co (fcc)	1645	$1120, 1350$	1311	1080	$1388 - 1390$
Ni (fcc)	397	$320, 820$	350	510	$624 - 631$
Gd (hcp)	334				293

While we quoted here results for "simple" metals, it is nowadays possible to investigate in the same manner the temperature dependent properties of complex multicomponent systems, e.g. half-metallic Heusler alloys [96] or dilute magnetic semiconductors [97]. In this way, materials for modern spintronic applications can be studied at physically relevant temperatures and their detailed magnetic properties can be predicted on the basis of quantum-mechanics. The combination of advanced numerical techniques and massively parallel supercomputers makes computational material science one of the most rapidly growing fields of physics with relevance for basic and applied science.

6 Examples: Low-dimensional magnets at surfaces

6.1 Non-collinear configurations of $3d$-impurities on ferromagnetic surfaces

In this section we provide some examples [98, 99, 100, 101] for non-collinear configurations of $3d$-dimers and multimers on the surfaces of ferromagnets. Here the unperturbed surfaces are ferromagnetic with a collinear moment configuration. Thus the Green function G^0 and t-matrix t^0 of the ideal surface are diagonal is spin-space

$$G^0(E) = \left\{ \begin{array}{cc} G^0_{\uparrow\uparrow}(E) & 0 \\ 0 & G^0_{\downarrow\downarrow}(E) \end{array} \right\} \quad t^0(E) = \left\{ \begin{array}{cc} t^0_{\uparrow\uparrow} & 0 \\ 0 & t^0_{\downarrow\downarrow} \end{array} \right\} \tag{59}$$

while non-collinear states lead to non-diagonal t-matrices for the impurity atoms and the surrounding substrate neighbors.

$$t(E) = \left\{ \begin{array}{cc} t_{\uparrow\uparrow} & t_{\uparrow\downarrow} \\ t_{\downarrow\uparrow} & t_{\downarrow\downarrow} \end{array} \right\} \tag{60}$$

The basic approximation with respect to non-collinearity consists of the assumption, that the exchange-correlation potential of each atom $\mathbf{R}^n$ has a unique quantisation axis $\mathbf{e}^n$, being common to the whole cell n and determined by the direction of the local moment $\mathbf{M}^n$ in cell n. In this local reference frame, the t-matrix t^n is diagonal

$$t^n_{loc} = \left\{ \begin{array}{cc} t^n_{\uparrow\uparrow} & 0 \\ 0 & t^n_{\downarrow\downarrow} \end{array} \right\} \tag{61}$$

and the local radial functions R^n_ℓ and H^n_ℓ are spin dependent as in a collinear calculation. However the Dyson equation describing the multiple scattering events has to be evaluated in a common global frame of reference, as e.g. determined by the magnetisation direction of the substrate. The corresponding transformed t^n-matrices are given by

$$t^n_{glob}(E) = U_n \, t^n_{loc}(E) \, U^t_n \tag{62}$$

where the rotation matrix U_n of spin space is given by

$$U_n = \left(\begin{array}{cc} \cos(\frac{\theta_n}{2}) \, e^{-i\phi_n/2} & -\sin(\frac{\theta_n}{2}) \, e^{-i\phi_n/2} \\ \sin(\frac{\theta_n}{2}) \, e^{i\phi_n/2} & \cos(\frac{\theta_n}{2}) \, e^{i\phi_n/2} \end{array} \right) \tag{63}$$

Here θ_n and ϕ_n are the polar angles defining the direction of the local moment $\mathbf{M}^n$ with respect to the substrate moments.

The basic reason for non-collinear states is "frustration", arising from the competition between ferromagnetic and antiferromagnetic coupling. In addition, also spin-orbit coupling can lead to a non-alignment of the local moments. However this is a very weak effect for transition metals, for which frustration is much more important. We will illustrate this in the following for transition metal dimers on the Ni(001) surface.

Let us start with single $3d$ adsorbate atoms on Ni(001). The calculations show, that the $3d$ adatoms have large and stable local moments. The moments of the Co, Fe and Mn adatoms couple ferromagnetically to the substrate moments, while the V and Cr moments prefer an antiferromagnetic coupling to the substrate. The situation of two $3d$-adatoms forming a dimer is illustrated in Fig. 20. Three kind of dimers are shown: Dimer 1 with the adatoms on nearest neighbor sites, Dimer 2 with the adatoms on second neighbor sites and Dimer 3 on fourth neighbor site. For the Dimer 2 and dimers with larger separation the interaction of the dimer atoms is very small and the configuration is dominated by the interaction with the substrate, meaning that these dimers show the same behavior as the single adatoms, coupling antiferromagnetically to the substrate in the case of V and Cr and ferromagnetically in the case of Mn and Fe, such that both adatoms are parallel aligned to each other. The same is also correct for the NN dimers of Fe or V, where the dimer atom interaction is strongly ferromagnetic (for Fe) or weakly antiferromagnetic (for V). In the case of the Cr and Mn dimers the situation is more complicated, since the interaction of the dimer atoms is strongly antiferromagnetic, favouring an antiferromagnetic pairing of the two moments. However, this is in contradiction to the interaction with the substrate moments, which as explained above, favours a parallel alignment of the impurity moments. Therefore frustration occurs, which can lead to a non-collinear ground state.

The situation is most easily explained, if a model operator in form of the classical Heisenberg model (2) applied to the interaction of the two adatoms $A = 1, 2$ and their interaction with the Ni moments, which for simplicity are assumed to be fixed, the Hamiltonian is

$$H = -J_{A-A} \cos(\theta_1 - \theta_2) - 4J_{A-Ni}(\cos\theta_1 + \cos\theta_2) \tag{64}$$

where θ_1 and θ_2 are the angles with the respect to the substrate magnetisation.

Let us now consider two typical spin configurations, shown in Fig. 21(a) and Fig. 21(b). Fig. 21(a) refers to a collinear configuration, which we call ferrimagnetic, since the two moments, being

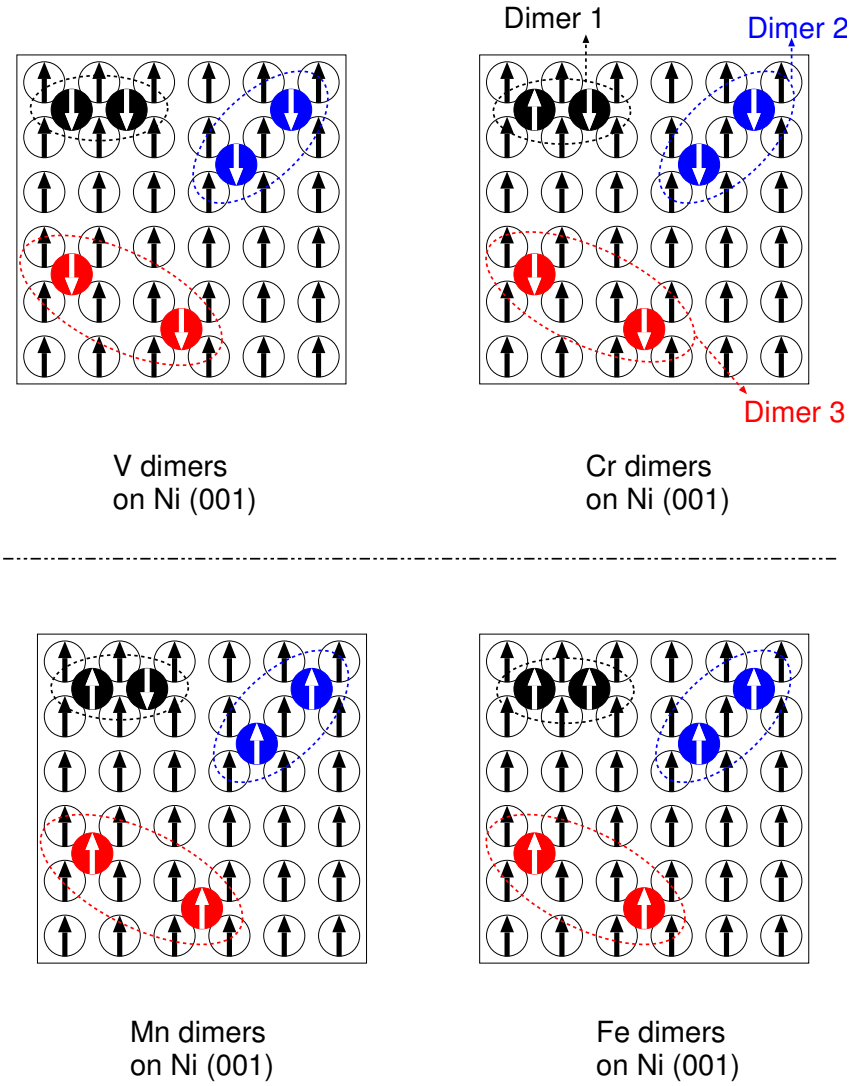
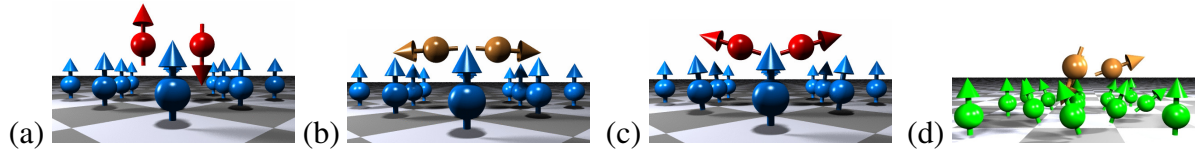


Fig. 20: *Different geometrical configurations considered for dimers at the surface of Ni(001). Dimer-1–type corresponds to the case where the atoms are first neighboring atoms, dimer-2–type where the atoms are second NN and finally dimer-3–type to fourth NN. The collinear magnetic ground state are also shown for V, Cr, Mn and Fe dimers.*

antiferromagnetically aligned, are not equivalent anymore, resulting in a small, but finite total moment. This configuration is also a selfconsistent solution of the Kohn-Sham equations, if the collinear constraint is removed. This can be understood for example from the Heisenberg model Eq. (64), since $\theta_1 = 0, \theta_2 = 180^o$ and small variations around these angles change the cos-values in Eq. (64) only in second order, so that the total energy is an extremum. The configuration in Fig.21(b) is noncollinear, but has the same energy as the collinear configuration (a), since in configuration (a) the interaction of the two adatoms with the substrate atoms cancel each other, while in configuration 21(b) they vanish for both atoms since $\cos\theta_1 = 0 = \cos\theta_2$. However this configuration is not a selfconsistent solution of the non-collinear Kohn-Sham equations, since a small variation $\Delta\theta_1$ and $\Delta\theta_2$ around the values of 90^o, respectively 270^o, changes the energy linearly in $\Delta\theta_1$ and $\Delta\theta_2$. Thus there exists a force which tilts the moments slightly towards or



(e)

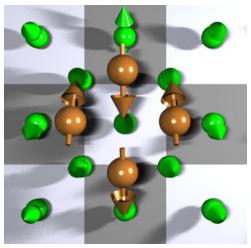


Fig. 21. *Most stable configurations of Cr and Mn dimer on Ni (in blue) obtained with (a) the collinear and (b)-(c) the non-collinear KKR method [98]. For Cr(Mn) the rotation angle with respect to the z axis is equal to 94.2°(72.6°) and the collinear(non-collinear) state is the ground state (see text). (d) shows a side view of the stable Cr trimer on the fcc Fe(001) surface (in green), with two Cr atoms pointing down (the second one cannot be seen), one Cr atom pointing up. (e) shows the stable Cr tetramer on the same fcc Fe surface.*

away from the surface, depending on the sign of J_{A-Ni}. In fact the configuration (b) is the non-collinear solution for a Cr-dimer. With a rotation angle of 94^o, deviating only slightly from 90^o (which can hardly be seen in the figure), a small energy is gained due to the antiferromagnetic coupling with the substrate $(J_{Cr-Ni} < 0)$. In contrast to this the configuration in Fig. 21(c) is the selfconsistent solution for a Mn-dimer, which prefers a ferromagnetic coupling with the substrate atoms ($J_{Mn-Ni} > 0$). Here the angle with respect to the z-axis is 73^o, the deviation from 90 % is much larger. The *ab initio* calculation shows, that this is the ground state for the Mn-dimer. However for the Cr-dimer the collinear solution of Fig. 21(a) is the ground state, which is in contradiction to the Heisenberg model and arises from small changes of the local moments upon rotation, an effect which cannot be described by this model.
In Fig. 21(d) and (e) we show two other non-collinear configurations obtained in the *ab initio* calculations, the configurations for compact Cr-trimers and Cr tetramers on fcc Fe_{3ML}/Cu(001). The exchange interactions are in this case very similar, except that the antiferromagnetic coupling of the Cr-adatoms to the Fe substrate atoms is considerably stronger. In both cases the Cr-Cr interaction is strongly antiferromagnetic. For the trimer, it is most important, that the effective interaction with the substrate moments is non-zero in the collinear configuration, but zero in the planar configuration. Thus the (basically) collinear configuration with the outer Cr-atoms antiferromagnetically aligned to the surface moments and the central Cr atom ferromagnetically aligned is favoured. However an additional small tilting occurs, in particular for the wrongly aligned central Cr atom, which further lowers the energy, so that also this configuration becomes non-collinear.
For the tetramer, the neighboring Cr-atoms couple again antiferromagnetically. For the in-plane configuration, similar to the dimer, the effective interaction with the substrate moments vanishes, however slight tiltings of the moments towards the surface lead to an additional energy gain stabilising the in-plane configuration.

6.2 Monolayers with complex spin structures

Antiferromagnetic interactions on a triangular lattice are the origin of frustrated spin systems. In recent years the epitaxial growth of such ultra-thin films has been studied intensively by various experimental techniques. In particular, pseudo-hexagonal c(8×2)Mn films on Cu(100) [102], Mn films on the (111) surfaces of fcc Pd [103], Ir [104], Cu [105, 106, 107], and MgO [108] and on the (0001) surface of Ru [109] and Co [110] have been prepared and analyzed. But also

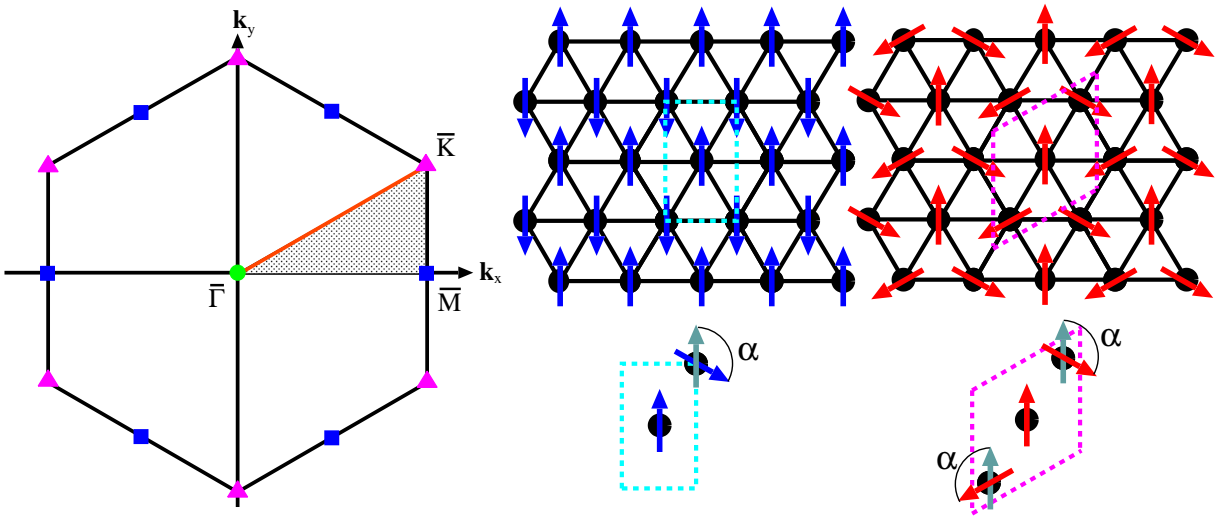


Fig. 22: *(Left:) The hexagon shows the first BZ of the 2D hexagonal Bravais lattice. The gray-shaded area indicates the irreducible part. (Center:) The RW-AFM structure. (Right:) the coplanar non-collinear Néel* ($120°$) *structure. Indicated are the corresponding two- and three-atom unit cells and the continous paths, which connect the corresponding magnetic structure to the FM state.*

other ultra-thin hexagonal films, e.g. Cr and V on Pt(111) and Ru(0001) [111, 112, 113], have been investigated.

To obtain an overview of all relevant spin-structures we develop first a zero-temperature phase diagram in the context of the Heisenberg model. As discussed in Section 2.1 the magnetic ground states are SSDWs, most likely with a commensurate propagation vector $\mathbf{q}_{\|}$ located at the high-symmetry points in the first 2DBZ of a 2D Bravais lattice. For the 2DBZ of the triangular (hexagonal) lattice, displayed in Fig. 22 (Left), the high-symmetry points are the corner points $\overline{\Gamma}$, $\overline{\mathrm{K}}$, and $\overline{\mathrm{M}}$ of the irreducible wedge of the 2DBZ (I2DBZ). The $\overline{\Gamma}$-point corresponds to the ferromagnetic solution. The $\overline{\mathrm{K}}$-point corresponds to a $120°$ Néel state (Fig. 22 (Center)), a 2D coplanar spin structure with three atoms in a $(\sqrt{3} \times \sqrt{3})$ R$30°$ unit cell for which the relative angle between the spins at the different sites is always $120°$. The $\overline{\mathrm{M}}$-point corresponds to row-wise antiferromagnetic (RW-AFM) configuration (Fig. 22 (Right)), which can be described by a rectangular unit cell with two antiferromagnetically aligned atoms. Magnetic ground states with incommensurate $\mathbf{q}_{\|}$-vectors are also possible preferentially with $\mathbf{q}_{\|}$-vectors from the connecting high-symmetry lines $\overline{\mathrm{M}}$-$\overline{\Gamma}$-$\overline{\mathrm{K}}$-$\overline{\mathrm{M}}$.

Along the line $\overline{\mathrm{M}}$-$\overline{\Gamma}$-$\overline{\mathrm{K}}$-$\overline{\mathrm{M}}$ we investigated the energetics within the Heisenberg model up the second nearest-neighbor interaction, i.e. including the exchange constants J_1, J_2. The results are summarized in Fig. 18 in terms of a zero-temperature phase diagram. Depending on the signs and values of J_1, and J_2 four kinds of possible magnetic ground states exist: FM, RW-AFM, $120°$, and the SSDW. If J_2 is zero or positive (ferromagnetic) than there are only two possible magnetic ground states, determined by the sign of J_1, the FM and the Néel state. But small values of J_2 are already sufficient to change the magnetic ground state and an infinite number of magnetic states becomes possible, the RW-AFM state or the incommensurable SSDW at any possible wave-vector $\mathbf{q}_{\|}$ at the high-symmetry line $\overline{\Gamma}$-$\overline{\mathrm{M}}$. Extending the model by including also J_3, a magnetic state with a $\mathbf{q}_{\|}$ at any high-symmetry line can become ground state.

Since the J's are rapidly varying functions of the number of d electrons, *ab initio* calculations

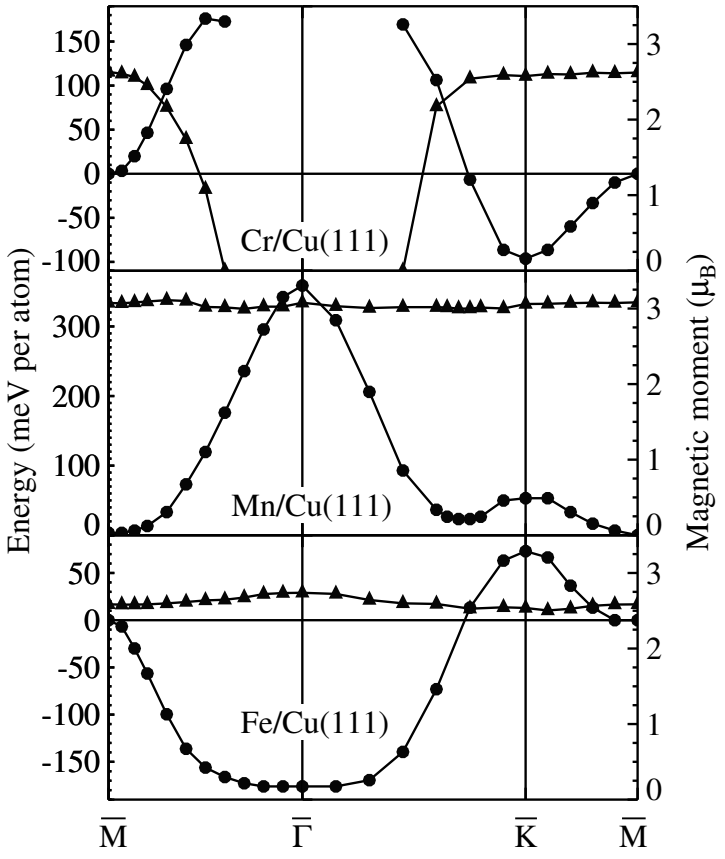


Fig. 23. *Calculated total energies (circles, left scale) and magnetic moments (triangles, right scale) for spin-spiral states in $3d$-UMLs with the Cu(111) geometry as function of the 2D wave vector $\mathbf{Q}_\parallel$ along the high symmetry lines of the 2DBZ. The energy is shown relative to the energy of the RW-AFM state.*

are carried out to determine the element specific ground states. Since the calculations are very time comsuming, the full overview has been worked out only for unsupported, free-standing monolayers (UML). Fig. 23 shows for the UMLs with the Cu lattice constant the total energy $E(\mathbf{Q}_\parallel)$ and the magnetic moments $M(\mathbf{Q}_\parallel)$ calculated for a discrete set of the spin-spiral $\mathbf{Q}_\parallel$ vectors along the high-symmetry lines. Among all the SSDWs calculated, the high-symmetry points have the lowest energies: the 120° Néel state ($\overline{\mathrm{K}}$-point) for Cr(111), the RW-AFM state ($\overline{\mathrm{M}}$-point) for Mn(111), and the FM state ($\overline{\Gamma}$-point) for Fe(111). For Fe and Mn, the $M(\mathbf{Q}_\parallel)$ are nearly a constant, but the Cr moments change drastically, as no ferromagnetic solution could be found for Cr(111). One more interesting observation is the local minimum of $E(\mathbf{Q}_\parallel)$ for Mn on the line $\overline{\Gamma}$-$\overline{\mathrm{K}}$, which is only 21 meV higher in energy than the RW-AFM state. We expect that a small change in the d-band filling, e.g. due to alloying with Fe, may change the energetics and an incommensurate SSDW may become the magnetic ground state.

For Mn, the lowest energy magnetic state found so far is the RW-AFM state, which corresponds to the commensurate SSDW state with one single $\mathbf{Q}_\parallel$-vector at the $\overline{\mathrm{M}}$-point of the 2DBZ, and the RW-AFM is also called single-$\mathbf{Q}_\parallel$ (1Q) state. In the 2DBZ there are three $\overline{\mathrm{M}}$-points corresponding to the three possible directions of the long axis of the RW-AFM unit cell on a triangular lattice. They are equivalent in symmetry, but are different to each other with $\mathbf{Q}_\parallel$-vectors, $\mathbf{Q}_\parallel^{(k)}$, for $k = 1, 2, 3$. Within the Heisenberg model the energy of each SSDW denoted by one of the three wave vectors $\mathbf{Q}_\parallel^{(k)}$ or any SSDW being an orthogonalized linear combination of those are degenerate. Higher order spin interactions (3) and (5) may lift this degeneracy and a so-called triple-$\mathbf{Q}_\parallel$ (3Q)-state, may become lower in energy. The 3Q-state is a three-dimensional non-collinear spin-structure on a 2D lattice (see Fig. 24) with four chemically identical atoms per surface unit-cell, where the relative angle between all nearest-neighbor spins is given by

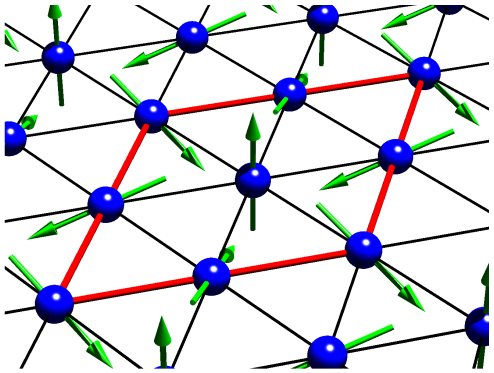


Fig. 24. *An image of the magnetic 3Q-structure, with spins pointing in all three directions of the spin-space. Note that, due to the neglect of the spin-orbit interaction only the relative orientation of the moments is specified.*

the tetrahedron angle of $109.47°$. The 3Q-state is formed as a linear combination of the three RW-AFM (1Q) structures orthogonal in spin-space, each having one of the three $\mathbf{Q}_{\parallel}^{(k)}$-vectors of the $\overline{\mathrm{M}}$-points:

$$\mathbf{m}(\mathbf{r}+\mathbf{R}_i) = m(\mathbf{r}) \times \frac{1}{\sqrt{3}} \sum_{k=1}^{3} e^{i\mathbf{Q}_{\parallel}^{(k)}\,\mathbf{R}_i}\,\hat{\mathbf{e}}^{(k)}, \tag{65}$$

where the $\hat{\mathbf{e}}^{(k)}$ are orthogonal unit vectors in spin space. We see that in the nearest-neigbhor approximation to the higher order exchange contributions the sign of K_1 and B_1 determine the sign of the energy difference $\Delta E = E_{3Q} - E_{1Q} = 16/3S^4(2K_1 + B_1)$ and thus whether the 3Q or the 1Q state becomes the magnetic ground state. From the *ab initio* calculations for the Mn UML in the geometry of Cu(111) we [53] found that the 3Q-state is 15 meV/atom lower in energy than the 1Q-state.
Calculations including the Cu(111) substrate show that the energy differences between different magnetic states change due to the present of the substrate, but the magnetic ground state remains unaltered: Cr/Cu(111) exhibits the 120° Néel state (2.35 μ_{B}), Mn/Cu(111) the 3Q-structure (2.74 μ_{B}), which is 17 meV lower in energy than the 1Q-state (3.00 μ_{B}), and Fe/Cu(111) is ferromagnetic (2.63 μ_{B}). On the Ag(111) substrate [114] the overall picture is the same, but two differences were noticed: V/Ag(111) is magnetic (2.19 μ_{B}) and exhibits as Cr/Ag(111) (3.65 μ_{B}) the 120° Néel state and the magnetic gound state of Mn/Ag(111) is the RW-AFM state (3.91 μ_{B}) and not the 3Q-state (3.88 μ_{B}). Fe/Ag(111) is ferromagnetic (3.02 μ_{B}). We believe that the complex spin-structures presented here, can be resolved using the spin-polarized scanning tunneling microscope in the constant-current mode [114, 115].

6.3 Chiral domain walls in Fe/W(110)

In this section, we illustrate the relevance of the DM interaction for the formation of chiral domain walls in Fe double layer (DL) on W(110), an ultrathin Fe film consisting of two atomic layers grown on the W(110)-surface. We choose this system, since its magnetic structure has been studied extensively by spin-polarized STM, cf. e.g. [116, 117, 118]. The magnetic pattern of the Fe DL on W(110) consists of a regular sequence of out-of-plane magnetized domains separated by domain walls [119, 45, 120]. The spatial orientation of the corresponding domain walls is determined by the crystal lattice and hardly influenced by the mesoscopic shape of the sample: The walls are preferably oriented normal to the $[001]$-direction (i.e. the magnetization changes along $[001]$ and remains constant along $[1\bar{1}0]$). In the following, we want to investigate this effect on the basis of the previously introduced micromagnetic model (50). If the DM term is irrelevant, the magnetization in the domain walls tries to avoid the hard axis and the

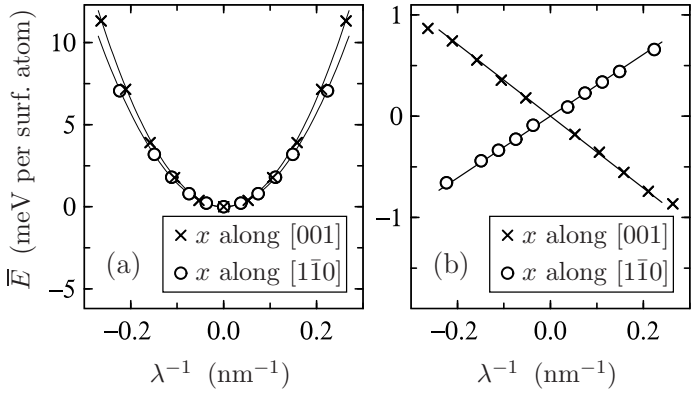


Fig. 25: *Energies of the homogeneous spin spirals obtained from electronic-structure calculations. The energies are plotted against the inverse of the period length λ. The results shown in (a) are obtained without spin-orbit interaction. Quadratic fits to these symmetric curves yield the spin stiffness A. (b) shows the odd part of $E(\lambda^{-1})$ obtained by including spin-orbit coupling. The slope of these curves at $\lambda^{-1} = 0$ correspond to D.*

magnetization rotation axis does not depend on the propagation direction. In this case, the energies $\sqrt{A\,K}$ of walls that are oriented in different crystallographic directions differ due to the spin stiffness, i.e. the value of A depends on the propagation direction. However, the values obtained for A by electronic-structure calculations do not change much for different propagations direction (cf. Fig. 25 left panel). A further indication of the relevance of the DM interaction in the studied system is the fact, that all domain walls that are observed within one sample show the same rotational direction [45, 120], this cannot be explained on the basis of symmetric exchange interactions. In the following we take the DM term into account, but we restrict our investigations to planar domain walls that can be described with the equations (50). If the propagation direction and the spin-rotation axis are both oriented along a high-symmetry line, we have to consider eight different walls that depend on six parameters. The walls are listed in Table 4 and their energies are given by equation $E = 4\sqrt{A\,K} - \pi\,|D|$.

We estimate values for the model parameters by electronic-structure calculations. A and D are obtained by using the approach introduced in Section 4.6. The results of these calculations are shown in Fig. 25. The anisotropy constant K consists of two main contributions, a term due to the anisotropic electronic energy and a term due to the magnetostatic interactions. The first term can be estimated directly from the electronic energies of collinear configurations with different spin quantization axes, the second term can be estimated from the summation of magnetic (dipole) moments Eq. (6). This sum converges fast, since we are considering a

	K_{001}		$K_{1\bar{1}0}$		
walls normal to [001] A_{001}	$+D_{001}$	$-D_{001}$	$D=0$	$D=0$	
walls normal to [1$\bar{1}$0] $A_{1\bar{1}0}$	$D=0$	$D=0$	$+D_{1\bar{1}0}$	$-D_{1\bar{1}0}$	[1$\bar{1}$0] [001]

Table 4: *Planar rotation path of domain-wall moments between two magnetic domains, with the magnetization pointing out-of-plane (⊗) or into-the plane (⊙), and corresponding model parameters.*

two-dimensional ultrathin magnetic film and not a magnetic bulk system. The results our calculations are summarized in Table 5 (for further details cf. Ref. [46]). The values given for the spin-stiffness constants A are fairly accurate, since they are obtained from an unambiguous fitting procedure on a curve on a large energy scale (cf. Fig. 25). The values given for D are less accurate, but they tell us the order of magnitude of the Dzyaloshinskii vector. The accuracy of the anisotropy constants K is not satisfactory, these calculations reach the limit of our computational method and we cannot rule out an error of a few $\mathrm{meV\,nm^{-1}}$ (note that $1\ \mathrm{meV\,nm^{-2}} \mathrel{\widehat{=}} 0.035$ meV per Fe atom).

		001	$1\bar{1}0$
spin stiffness	$A\ /\ (\,\mathrm{meV}\,)$	58.8	51.1
DM interaction	$D\ /\ (\,\mathrm{meV\,nm^{-1}}\,)$	-8.0	6.9
anisotropy energy	$K\ /\ (\,\mathrm{meV\,nm^{-2}}\,)$	1.4	3.0

Table 5: *Theoretically predicted model parameters converted into areal densities. The crystallographic directions refer to the indices used in Table 4.*

Inserting the values given in Table 5 in equation $E = 4\sqrt{A\,K} - \pi\,|D|$ yields the wall energies

$$\begin{aligned} 4\sqrt{A_{001}\,K_{001}} - \pi\,|D_{001}| &= 11\ \mathrm{meV\,nm^{-1}}\ , \\ 4\sqrt{A_{001}\,K_{1\bar{1}0}} &= 53\ \mathrm{meV\,nm^{-1}}\ , \\ 4\sqrt{A_{1\bar{1}0}\,K_{001}} &= 34\ \mathrm{meV\,nm^{-1}}\ , \\ 4\sqrt{A_{1\bar{1}0}\,K_{1\bar{1}0}} - \pi\,|D_{1\bar{1}0}| &= 28\ \mathrm{meV\,nm^{-1}}\ . \end{aligned} \tag{66}$$

We find the observed wall orientation cannot be explained by the spin stiffness alone, and the DM interaction is strong enough to compete with the other quantities. According to the values given in Table 5 the lowest energy indeed corresponds to a wall oriented normal to the $[001]$ direction cf. Fig. 4 as observed experimentally. The negative sign of D_{001} corresponds to a right-rotating wall. The formation of a domain wall does not lower the energy of the system. This implies that the collinear state is the ground state. In agreement with SP-STM experiments we find walls with a specific rotational direction that are oriented normal to the $[001]$ direction. We identify these walls as right-rotating Néel-type walls. Based on the microscopic understanding we had been able to solve a longstanding open problem of why the domain orientation is sensitive to the underlying crystal structure rather than to the surface geometry. The studied system nicely illustrates the relevance of the Dzyaloshinskii-Moriya interactions for magnetic surfaces. Chiral domain walls are an important new type of domain-wall with interesting new properties relevant to concepts in information storage such as the racetrack memory concept.

Acknowledgements

I gladly acknowledge the discussions with Dr. Gustav Bihlmayer and Dr. Samir Lounis.

References

[1] M. Bode, M. Heide, K. von Bergmann, P. Ferriani, S. Heinze, G. Bihlmayer, A. Kubetzka, O. Pietzsch, S. Blügel, and R. Wiesendanger, Nature **447**, 190 (2007).

[2] J. Sticht, K. H. Höck, and J. Kübler, J. Phys.: Condens. Matter **1**, 8155 (1989).

[3] J. Kübler, *Theory of Itinerant Magnetism*, Oxford University Press (2000).

[4] L. M. Sandratskii and J. Kübler, Newsletter 14 (1996), http://www.psi-k.org/newsletters/-News_14/Highlight_14.pdf .

[5] L. M. Sandratskii, Adv. in Phys. **47**, 91 (1998).

[6] Ph. Kurz, F. Förster, G. Bihlmayer, S. Blügel, and L. Nordström, Newsletter **38** (2000), http://www.psi-k.org/newsletters/News_38/newsletter_38.pdf .

[7] Ph. Kurz, F. Förster, L. Nordström, G. Bihlmayer, and S. Blügel, Phys. Rev. B **69**, 024415 (2004).

[8] F. Schiller, D. V. Vyalikh, V. D. P. Servedio, and S. L. Molodtsov, Phys. Rev. B **70**, 174444 (2004).

[9] R. Hafner, D. Spisak, R. Lorenz and J. Hafner, Phys. Rev. B **65**, 184432 (2002).

[10] K. Hirai, J. Phys. Soc. Jpn. **62**, 690 (1993).

[11] E.O. Wollan and W.C. Koehler, Phys. Rev. **100**, 545 (1955).

[12] F. Keffer, Encyclopedia of Physics, Vol. 18/2, (1966) edited by H. P. J. Wijn (Berlin: Springer).

[13] J. M. D. Coey, Can. J. Phys. **65**, 1210, (1987).

[14] S. Geller, J. Chem. Phys. **24** 1236 (1956).

[15] T. Kimura, S. Ishihara, H. Shintani, T. Arima, K. T. Takahashi, K. Ishizaka, and Y. Tokura, Phys. Rev. B **68**, 060403 (2003).

[16] C. Zener, Phys.Rev. **82**, 403 (1951).

[17] C. Zener, Phys.Rev. **81**, 440 (1951).

[18] B. L. Gyorffy, A. J. Pindor, J. Staunton, G. M. Stocks and H. Winter, J. Phys. F **15**, 1337 (1985).

[19] V. P. Antropov, N. N. Harmon and A. N. Smirnov, J. Magn. Magn. Mater. **200**, 148 (1999).

[20] I. Dzyaloshinsky, J. Phys. Chem. Solids **4**, 241 (1958).

[21] T. Moriya, Phys. Rev. **120**, 91 (1960).

[22] A. N. Bogdanov, D. A.Yablonskii, Sov. Phys. JETP **68**, 101 (1989).

[23] A. Bogdanov and A. Hubert, J. Magn. Magn. Mater. **138**, 255 (1994).

[24] T.H.R. Skyrme, Nucl. Phys. **31**, 556 (1962).

[25] A. Hubert, R. Schäfer, *Magnetic Domains* (Springer-Verlag, Berlin, 1998).

[26] A. Fert, V. Cros and J. Sampaio, Nature Nanotech. **8**, 152 (2013).

[27] T. Schulz, R. Ritz, A. Bauer, M. Halder, M. Wagner, C. Franz, C. Pfleiderer, K. Everschor, M. Garst, A. Rosch, Nature Physics **8**, 301 (2012).

[28] Y. Aharonov and A. Stern, Phys. Rev. Lett. **69**, 3593 (1992).

[29] P. Bruno, Phys. Rev. Lett. **93**, 096806 (2004).

[30] S. Mühlbauer, B. Binz, F. Jonietz, C. Pfleiderer, A. Rosch, A. Neubauer, R. Georgii, P. Böni, Science **323**, 915 (2009).

[31] X. Z. Yu, Y. Onose, N. Kanazawa, J. H. Park, J. H. Han, Y. Matsui, N. Nagaosa, and Y. Tokura, Nature **465**, 901 (2010).

[32] C. Pappas, E. Lelièvre-Berna, P. Falus, P. M. Bentley, E. Moskvin, S. Grigoriev, P. Fouquet, B. Farago, Phys. Rev. Lett. **102**, 197202 (2009).

[33] N. Kanazawa, Y. Onose, T. Arima, D. Okuyama, K. Ohoyama, S. Wakimoto, K. Kakurai, S. Ishiwata, Y. Tokura, Phys. Rev. Lett. **106**, 156603 (2011).

[34] H. Wilhelm, M. Baenitz, M. Schmidt, U. K. Rößler, A. A. Leonov, A. N. Bogdanov, Phys. Rev. Lett. **107**, 127203 (2011).

[35] X. Z. Yu, N. Kanazawa, Y. Onose, K. Kimoto, W. Z. Zhang, S. Ishiwata, Y. Matsui, Y. Tokura, Nature Materials **10**, (2011).

[36] S.V. Grigoriev, N. M. Potapova, S.-A. Siegfried, V. A. Dyadkin, E.V. Moskvin, V. Dmitriev, D. Menzel, C. D. Dewhurst, D. Chernyshov, R. A. Sadykov, L. N. Fomicheva, A.V. Tsvyashchenko, Phys. Rev. Lett. **110**, 207201 (2013).

[37] K. Shibata, X. Z. Yu, T. Hara, D. Morikawa, N. Kanazawa, K. Kimoto, S. Ishiwata, Y. Matsui, Y. Tokura, Nature Nanotechnology **8**, 723 (2013).

[38] P.Bak and M.H.Jensen J.Phys.C: Solid State Phys., **13**, L881 (1980).

[39] P. Ferriani, K. von Bergmann, E.Y. Vedmedenko, S. Heinze, M. Bode, M. Heide, G. Bihlmayer, S. Blügel and R. Wiesendanger, Phys. Rev. Lett. **101**, 027201 (2008).

[40] K. von Bergmann, S. Heinze, M. Bode, E. Y. Vedmedenko, G. Bihlmayer, S. Blügel, R. Wiesendanger, Phys. Rev. Lett. **96**, 167203 (2006).

[41] K. von Bergmann, S. Heinze, M. Bode, G. Bihlmayer, S. Blügel, R. Wiesendanger, New J. Phys. **9**, 396 (2007).

[42] S. Heinze, K. von Bergmann, M. Menzel, J. Brede, A. Kubetzka, R. Wiesendanger, G. Bihlmayer, S. Blügel, Nature Physics **7**, 713-718 (2011).

[43] N. Romming, M. Menzel, C. Hanneken, J. E. Bickel, B. Wolter, K. von Bergmann, André Kubetzka, R. Wiesendanger, Science **341**, 636 (2013).

[44] Matthias Menzel, Yuriy Mokrousov, Robert Wieser, Jessica E. Bickel, Elena Vedmedenko, Stefan Blügel, Stefan Heinze, Kirsten von Bergmann, André Kubetzka, and Roland Wiesendanger Phys. Rev. Lett. **108**, 197204 (2012).

[45] A. Kubetzka, M. Bode, O. Pietzsch, and R. Wiesendanger, Phys. Rev. Lett. **88**, 057201 (2002).

[46] M. Heide, G. Bihlmayer, and S. Blügel Phys. Rev. B **78**, 140403(R) (2008).

[47] G. Chen, J. Zhu, A. Quesada, J. Li, A. T. NDiaye, Y. Huo, T. P. Ma, Y. Chen, H.Y. Kwon, C. Won, Z. Q. Qiu, A. K. Schmid, and Y. Z. Wu Phys. Rev. Lett. **110**, 177204 (2013).

[48] Kh. Zakeri, Y. Zhang, J. Prokop, T.-H. Chuang, N. Sakr, W. X. Tang, and J. Kirschner Phys. Rev. Lett. **104**, 137203 (2010).

[49] W. Heisenberg, Z. Phys. **49**, 619 (1928).

[50] M. Takahashi, J. Phys. C **10**, 1289 (1977).

[51] J.J. Sakurai, Modern Quantum Mechanics Revised Edition (Addison-Wesley, Reading, MA, 1994).

[52] K. Yoshida and S. Inagaki, J. Phys. Soc. Japan **50**, 3268 (1981).

[53] Ph. Kurz, G. Bihlmayer, K. Hirai and S. Blügel, Phys. Rev. Lett. **86** , 1106 (2001).

[54] S.-W. Cheong and M. Mostovoy, Mature Materials **6**, 13 (2007).

[55] D. A. Smith, J. Magn. Magn. Mater. **1**, 214 (1976).

[56] A. Fert and P. M. Levy, Phys. Rev. Lett. **44**, 1538 (1980).

[57] A. Fert Mater. Sci. Forum, **5960**, 439 (1990).

[58] M. Kataoka, O. Nakanishi, A. Yanase, and J. Kanamori, J. Phys. Soc. Jap. **53**, 3624 (1984).

[59] J. Bouaziz, M. dos Santos Dias, A. Ziane, M. Benakki, S. Blügel, S. Lounis, to be published (2014).

[60] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[61] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[62] U. von Barth and L. Hedin, J. Phys. C: Solid State Phys. **5**, 1629 (1972).

[63] S. Di Napoli, A. M. Llois, G. Bihlmayer, S. Blügel, M. Alouani, and H. Dreysse, Phys. Rev. B **70** 174418 (2004).

[64] I. Turek, J. Kudrnovský, M. Diviš, P. Franek, G. Bihlmayer, and S. Blügel, Phys. Rev. B **68**, 224431 (2003).

[65] D. J. Singh and J. Ashkenazi, Phys. Rev. B **46**, 11570 (1992).

[66] V. L. Moruzzi, J. F. Janak, and A. R. Williams, *Calculated Electronic Properties of Metals*, (Pergamon, New York, 1978).

[67] S. Shallcross, A. E. Kissavos, V. Meded, and A. V. Ruban, Phys. Rev. B **72**, 104437 (2005).

[68] Ph. Kurz, G. Bihlmayer, and S. Blügel, J. Phys.: Condens. Matter **14**, 6353 (2002).

[69] H. W. White, B. J. Beaudry, P. Burgardt, S. Legvold, and B. N. Harmon, AIP Conf. Proc. **29**, 329 (1975).

[70] G. Vignale and M. Rasolt, Phys. Rev. B **37**, 10685 (1988).

[71] H. A. Bethe and E. E. Salpeter *Quantum Mechanics of One- and Two-Electron Systems* (New York, Plenum 1977).

[72] T. Thonhauser, D. Ceresoli, D. Vanderbilt, and R. Resta, Phys. Rev. Lett. **95**, 137205 (2005).

[73] I. Yang, S. Y. Savrasov, and G. Kotliar, Phys. Rev. Lett. **87**, 216405 (2001).

[74] J. Stöhr, J. Magn. Magn. Mater. **200**, 470 (1999).

[75] G. Bihlmayer in: *Magnetism goes Nano, Schriften des FZ-Jülich: Matter and Materials (26)* (Jülich, 2005).

[76] M. Weinert, R. E. Watson, and J. W. Davenport, Phys. Rev. B **32**, 2115 (1985).

[77] V. P. Antropov, M. I. Katsnelson, M. van Schilfgaarde, and B. N. Harmon, Phys. Rev. Lett. **75**, 729 (1995).

[78] V. P. Antropov, M. I. Katsnelson, B. N. Harmon, M. van Schilfgaarde, and D. Kusnezov, Phys. Rev. B **54**, 1019 (1996).

[79] G. M. Stocks, B. Újfalussy, X.-D. Wang, Y. Wang, D. M. C. Nicholson, W. A. Shelton, A. Canning, and B. L. Györffy, Phil. Mag. B **78**, 665 (1998).

[80] P. H. Dederichs, S. Blügel, R. Zeller, and H. Akai, Phys. Rev. Lett. **53**, 2512 (1984).

[81] B. Újfalussy, X.-D. Wang, D. M. C. Nicholson, W. A. Shelton, G. M. Stocks, Y. Wang, and B. L. Györffy, J. Appl. Phys. **85**, 4824 (1999).

[82] P. W. Anderson, *Theory of Magnetic Exchange Interactions: Exchange in Insulators and Semiconductors*, Solid State Physics **14**, 99 (1963).

[83] K. Yosida, *Theory of Magnetism* (Spinger, Berlin-Heidelberg, 1996).

[84] C. Herring, in *Magnetism*, (Ed. G. Rado and H. Suhl) (Academic, New York, 1966).

[85] L. M. Sandratskii, Phys. Status Solidi B **136**, 167 (1986).

[86] L. M. Sandratskii, J. Phys.: Condens. Matter **3**, 8565 (1991).

[87] N. Ashcroft and N. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).

[88] M. Ležaić, P. Mavropoulos, G. Bihlmayer, and S. Blügel, Phys. Rev. B **88**, 134403 (2013).

[89] A. I. Liechtenstein, M. I. Katsnelson, V. P. Antropov, and V. A. Gubanov, J. Magn. Magn. Mater. **67**, 65 (1987).

[90] A. R. Mackintosh and O. K. Andersen, *Electrons at the Fermi Surface*, page 149, ed. M. Springford, Cambridge Univ. Press, London (1980).

[91] M. Heide, G. Bihlmayer, and S. Blügel, Physica B **404**, 2678 (2009).

[92] N. M. Rosengaard and Börje Johansson, Phys. Rev. B **55**, 14975 (1997).

[93] M. Pajda, J. Kudrnovský, I. Turek, V. Drchal, and P. Bruno, Phys. Rev. B **64**, 174402 (2001).

[94] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[95] I. Turek, J. Kudrnovský, G. Bihlmayer, and S. Blügel, J. Phys.: Condens. Matter **15**, 2771 (2003).

[96] E. Şaşıoğlu, L. M. Sandratskii, and P. Bruno, J. Appl. Phys. **98**, 063523 (2005).

[97] K. Sato, P. H. Dederichs, and H. Katayama-Yoshida, J. Supercond. **18**, 33 (2005).

[98] S. Lounis, Ph. Mavropoulos, P. H. Dederichs and S. Blügel, Phys. Rev. B **72**, 224437 (2005).

[99] S. Lounis, Ph. Mavropoulos, R. Zeller, P. H. Dederichs and S. Blügel
Phys. Rev. B **75**, 174436 (2007).

[100] S. Lounis, M. Reif, Ph. Mavropoulos, L. Glaser, P. H. Dederichs, M. Martins, S. Blügel and W. Wurth, Eur. Phys. Lett. **81**, 47004 (2008).

[101] S. Lounis, P. H. Dederichs and S. Blügel, Phys. Rev. Lett. **101**, 107204 (2008).

[102] T. Flores, M. Hansen, and M. Wuttig, Surf. Sci. **279**, 251 (1992).

[103] D. Tian, H. Li, S. Wu, F. Jona, and P. Marcus, Phys. Rev. B **45**, 3749 (1992).

[104] S. Andrieu, H. Fischer, M.Piecuch, A. Traverse, and J. Mimault, Phys. Rev. B **54**, 2822 (1996).

[105] D. Tian, A. Begley, and F. Jona, Surf. Sci. Lett. **273**, 393 (1992).

[106] I. Grigorov and J. Walker, J. Appl. Phys. **81**, 3907 (1997).

[107] I. Grigorov, J. Walker, M. Hawley, G.B. and M.Lu tt,and M. Fitzsimmons, J. Appl. Phys. **83**, 7010 (1998).

[108] I. Grigorov, I.-L. Siu, M. Fitzsimmons, and J. Walker, Phys. Rev. Lett. **82**, 5309 (1999).

[109] A. Arrott, B. Heinrich, S. Purcell, J. Cochran, and L. Urquhart, J. Appl. Phys. **61**, 3721(1987).

[110] K. Ounadjela, P. Vennegues, Y. Henry, A. Michel, V. Pierron-Bohnes, and J. Arabski, Phys. Rev. B **49**, 8561 (1994).

[111] L. Zhang, M. Kuhn, and U. Diebold, Surf. Sci. **371**, 223 (1997).

[112] M. Albrecht, J. Pohl, H. Wider, E. Malang, J. Kohler, K. Friemelt, and E. Bucher, Surf. Sci. **397**, 354 (1998).

[113] M. Sambi and G. Granozzi, Surf. Sci. **426**, 235 (1999).

[114] S. Heinze, Ph. Kurz, D. Wortmann, G. Bihlmayer, and S. Blügel, Appl. Phys. A **75**, 25 (2002).

[115] D. Wortmann, S. Heinze, Ph. Kurz, G. Bihlmayer, and S. Blügel, Phys. Rev. Lett **86**, 4132 (2001).

[116] M. Pratzer, H. J. Elmers, M. Bode, O. Pietzsch, A. Kubetzka, and R. Wiesendanger, Phys. Rev. Lett. **87**, 127201 (2001).

[117] M. Bode, Rep. Prog. Phys. **66**, 523 (2003).

[118] E. Y. Vedmedenko, A. Kubetzka, K. von Bergmann, O. Pietzsch, M. Bode, J. Kirschner, H. P. Oepen, and R. Wiesendanger, Phys. Rev. Lett. **92**, 77207 (2004).

[119] M. Bode, O. Pietzsch, A. Kubetzka, S. Heinze, and R. Wiesendanger, Phys. Rev. Lett. **86**, 2142 (2001).

[120] A. Kubetzka, O. Pietzsch, M. Bode, and R. Wiesendanger, Phys. Rev. B **67**, 020401(R) (2003).

C 5 Function renormalization group for weakly correlated systems[1]

Carsten Honerkamp
Institute for Theoretical Solid State Physics
RWTH Aachen University

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercom-

1 Introduction

This spring school contribution intends to address and combine two different issues. The first issue a methodical one, how one can use the functional renormalization group (fRG) for exploring the possible low-temperature states of interacting electron lattice models. The second issue is how multi-band systems can show interesting additional aspects and new possibilities in their ground state properties, i.e. how novel order parameters become important and how usual order is altered by multi-band effects.

We use the term 'weakly correlated' electrons in the title in the following sense: First of all, our theoretical approach, the fRG, is in its present implementation, due to the approximations committed, certainly only valid up to a moderate strength of the electron-electron interaction on the lattice. Second, the two main material systems described in this article, the iron pnictide superconductors[1] and layered graphene[2], are known to be rather on the weak coupling side of the spectrum. Especially in bulk graphene most but not all experiments available today can be described theoretically without any explicit reference to electron-electron interactions. Nevertheless, as electrons on the honeycomb lattice of graphene have become an important new model system, it is important to know the weak coupling expectations for the ground state. Furthermore, recent experiments on bi- and trilayer systems show gap openings at low T that have been interpreted as being interaction effects[3, 4].

We will start out with a brief recapitulation of instabilities due to electron-electron interactions in many-fermion lattice systems. Then we will describe the functional RG formalism adapted for two-dimensional lattices as a tool to study the instability question in an unbiased and physically informative way. We also mention how this approach can be embedded in a more material-specific context. Then we give example for the one-band Hubbard model. In the second part of the article we first discuss how the picture is augmented in the many-band systems. For the iron arsenides, we will find an additional gap variation caused by multi-band effects. In multi-layer graphene, we will see that among the possible ordered ground states, also novel interaction-induced quantum spin-Hall states might be realized.

In this article we mainly highlight some works of the author and his colleagues. However, the fRG in rather similar form has been applied in many other groups as well, for a growing number of relevant problems. Many of these works are described and referenced in the recent reviews by Metzner et al.[5] and Platt et al.[6]. For a thorough introduction into the formalism in less compactified form, see for example the textbook by Kopietz and Schütz[7]. Parts of this manuscript have been adapted from Ref. [8]. Besides itinerant electron systems, there is also a related development where the fRG is used for the analysis for two-dimensional quantum spin systems[9].

2 Instabilities for interacting fermions

Here we show how a perturbative evaluation of the two-particle interaction vertex for electrons on lattices can lead to various infinities at low temperatures. The basic object of study is the one-particle irreducible (1PI) two-particle vertex, which can be constructed from the two-particle Green's function of the theory[10].

2.1 Model, functional integral formalism and Green's functions

First we introduce the 'standard model for correlated electrons', the one-band Hubbard model on the two-dimensional square lattice. In terms of fermionic annihilation and creation operators on the lattice sites i with spin variable s, the Hamiltonian is given by

$$H = -t \sum_{\substack{\langle i,j\rangle,\\ \text{n.n.}}} \left(c^\dagger_{i,s} c_{j,s} + c^\dagger_{j,s} c_{i,s} \right) \quad - \quad t' \sum_{\substack{\langle\langle i,j\rangle\rangle,\\ \text{n.n.n.}}} \left(c^\dagger_{i,s} c_{j,s} + c^\dagger_{j,s} c_{i,s} \right) + U \sum_i n_{i,\uparrow} n_{i,\downarrow} \,, \qquad (1)$$

with t and t' being the hopping amplitudes to nearest neighbors and diagonal next-nearest neighbors on the square lattice. U is the Hubbard on-site repulsion. Together with the chemical potential μ, the hopping parameters t and t' determine the shape of the Fermi surface. The band structure including μ is given by $\epsilon(\vec{k}) = -2t \cos k_x \cos k_y - 4t' \cos k_x \cos k_y - \mu$.
In a more realistic setting, the orbitals on the sites i should be understood as (maximally localized) Wannier states. The parameters t and t' can then be obtained from fitting the hopping dispersion to the full band structure, or from computing matrix elements with the non-interacting Hamiltonian, as explained in various textbooks. The Hubbard-U is, in the simplest approach, given by the onsite matrix elements of the Coulomb repulsion. In recent years, the role of the screening due to the other bands not considered in the Hubbard model has been recognized[11]. The following technical subsections are formulated for general one-band models, but we will return to the Hubbard model in the more concrete evaluations.
The grand-canonical partition function of the many-fermion problem of interest shall given by

$$Z = \mathrm{Tr}\, e^{-\beta H(c^\dagger_{i,s}, c_{i,s})} \,. \qquad (2)$$

For simplicity, we have absorbed the chemical potential into the Hamiltonian already.
In the functional integral formalism with anticommuting Grassmann fields ψ and $\bar{\psi}$, the partition function of the many-fermion system is rewritten in terms of integral over the exponential of the action S of the system,

$$Z = \int D[\psi, \bar{\psi}]\, e^{-S(\psi,\bar{\psi})} \,. \qquad (3)$$

The fields ψ and $\bar{\psi}$ are antiperiodic in imaginary time τ ranging in the interval from 0 to $\beta = 1/T$, the inverse temperature. In the Matsubara representation, one transforms the τ-dependence to fermionic Matsubara frequencies $\omega_n = i\pi T(2n-1)$ with n being integer. Then the fermionic fields depend on quantum numbers that can be collected in a multi-index $K = (\vec{k}, i\omega, s)$, with wave vector $\vec{k}$ and spin index s. For a usual many-fermion lattice system, the non-interacting quadratic part of the fermionic action then has the form

$$S_0 = \sum_{\vec{k},\omega_n,s} \left[-i\omega_n + \epsilon(\vec{k}) \right] \bar{\psi}(\vec{k}, \omega_n, s) \psi(\vec{k}, \omega_n, s) \,. \qquad (4)$$

Here, we have refocussed on a one-band description and $\epsilon(\vec{k})$ is the band energy of this band. To make the notation more simple it is often useful to write instead

$$S_0 = \sum_K Q(K) \bar{\psi}(K) \psi(K) = (\bar{\psi}, Q\psi) \,. \qquad (5)$$

with the quadratic part $Q(K) = \left[-i\omega_n + \epsilon(\vec{k}) \right]$. In the last equality we have introduced the short notation $(\bar{\phi}, \eta) = \sum_K \bar{\phi}(K)\eta(K)$.

The single-particle Green's function $G(K)$ is then given by

$$G(K) = -\langle \psi(K)\bar{\psi}(K)\rangle = -\frac{\int D[\psi,\bar{\psi}]\,\psi(K)\bar{\psi}(K)e^{-S(\psi,\bar{\psi})}}{\int D[\psi,\bar{\psi}]\,e^{-S(\psi,\bar{\psi})}}\,. \tag{6}$$

In the non-interacting case, we have

$$G_0(K) = -Q^{-1}(K) = \frac{1}{i\omega_n - \epsilon(\vec{k})}\,. \tag{7}$$

Higher-order correlation functions are defined as

$$\begin{aligned} G^{(2n)}(K_1,\dots K_n,K_1',\dots,K_n') &= (-1)^n\langle\psi(K_1)\dots\psi(K_n)\bar{\psi}(K_n')\dots\bar{\psi}'(K_1)\rangle \\ &= (-1)^n\frac{\int D[\psi,\bar{\psi}]\,\psi(K_1)\dots\psi(K_n)\bar{\psi}(K_n')\dots\bar{\psi}'(K_1)\,e^{-S(\psi,\bar{\psi})}}{\int D[\psi,\bar{\psi}]e^{-S(\psi,\bar{\psi})}}\,. \end{aligned}$$

It is now simple to see that these correlation functions can be obtained from a generating functional

$$\mathcal{G}(\eta,\bar{\eta}) = Z^{-1}\int D[\psi,\bar{\psi}]\,e^{-S(\psi,\bar{\psi})}e^{\sum_K\left[\bar{\eta}(K)\psi(K)+\bar{\psi}(K)\eta(K)\right]}\,, \tag{8}$$

by taking derivatives,

$$G^{(2n)}(K_1,\dots K_n,K_1',\dots,K_n') = \left.\frac{\delta^{2n}\mathcal{G}(\eta,\bar{\eta})}{\delta\bar{\eta}(K_1)\dots\delta\bar{\eta}(K_n)\delta\eta(K_n')\dots\delta\eta(K_1')}\right|_{\eta=\bar{\eta}=0}\,. \tag{9}$$

One can now use perturbation theory in the interactions V and finds that the functional $\mathcal{G}(\eta,\bar{\eta})$ generates non-connected diagrams as well connected diagrams. As we will see later, taking the logarithm removes these disconnected parts, and the perturbation expansion has less terms. The **generating functional for the connected correlation functions** is usually (there is some freedom regarding the normalization of the functional integral) defined as

$$\mathcal{G}^c(\eta,\bar{\eta}) = -\ln\left[Z\mathcal{G}(\eta,\bar{\eta})\right] = -\ln\left\{\int D[\psi,\bar{\psi}]\,e^{-S(\psi,\bar{\psi})}e^{\sum_K[\bar{\eta}(K)\psi(K)+\bar{\eta}(K)\psi(K)]}\right\}\,. \tag{10}$$

Expanding $\mathcal{G}^c(\eta,\bar{\eta})$ in the source fields leads to a formal power series with the connected Green functions as coefficients,

$$\begin{aligned} \mathcal{G}^c(\eta,\bar{\eta}) &= -\ln Z + (\bar{\eta},G_c^{(2)}\eta) \\ &+ \frac{1}{(2!)^2}\sum_{K_1,K_2,K_1',K_2'} G_c^{(4)}(K_1,K_2;K_1',K_2')\,\bar{\eta}(K_1)\bar{\eta}(K_2)\eta(K_2')\eta(K_1') + \dots\,. \end{aligned} \tag{11}$$

2.2 Perturbation expansion for the two-particle vertex

Next we can study some aspects of the effective interactions, i.e. the interactions that act between electrons not only due to the bare interaction in the Hamiltonian, but also due to multiple scattering events. This information is contained in the connected part of the two-particle correlation function. The connected part of this two-particle Green's function can be understood as interaction contribution as it is only nonzero for nonzero interactions. This part turns out to

be a product of four fully renormalized Green's functions $G^{(2)}$ times a one-particle irreducible vertex that includes all possible (1PI) multiple scattering events[10],

$$G_c^{(4)}(K_1, K_2; K_1', K_2') = \\ G^{(2)}(K_1)\, G^{(2)}(K_2)\, \Gamma^{(4)}(K_1, K_2; K_1', K_2') G^{(2)}(K_1')\, G^{(2)}(K_2')\ . \tag{12}$$

The two-particle Green's function $G_c^{(4)}$ or the interaction vertex $\Gamma^{(4)}$ are the basic objects in the computation of many observables. For example, spin and charge susceptibilities or conductivities can be readily expressed as expectation values over two barred and two unbarred fields. The connected part contains the difference of the interacting susceptibilities compared to the bare ones. Instabilities towards spontaneous symmetry breaking occur as divergences of the connected two-particle Green's function or the vertex. Furthermore, $\Gamma^{(4)}$ can be understood as effective scattering amplitude between the excitations specified by the quantum numbers $K_{1/2}$ and $K'_{1/2}$.

Our main emphasis will now be on the perturbation expansion for $G_c^{(4)}(K_1, K_2, K_1', K_2')$ or the corresponding vertex. This will lead us to various divergences at low temperatures which show that in many systems the Fermi liquid is ultimately not the stable ground state. This will then motivate the study of symmetry-broken states that replace the Fermi liquid as ground state.

Let us for simplicity choose a Hubbard-type interaction, translated from (1) into the Grassmann picture (with N as the number of lattice sites),

$$S_I = \frac{U}{N} \int_0^\beta d\tau \sum_{\vec{k},\vec{k}',\vec{q}} \bar{\psi}_\uparrow(\vec{k}+\vec{q},\tau)\bar{\psi}_\downarrow(\vec{k}'-\vec{q},\tau)\psi_\downarrow(\vec{k}',\tau)\psi_\uparrow(\vec{k},\tau)\,. \tag{13}$$

In Matsubara space this becomes

$$S_I = \frac{TU}{N} \sum_{k,k',q} \bar{\psi}_\uparrow(k+q)\bar{\psi}_\downarrow(k'-q)\psi_\downarrow(k')\psi_\uparrow(k)\,. \tag{14}$$

The perturbation expansion for the two-particle Green's function then reads, using $K = (p, s)$ and so on,

$$\begin{aligned} G^{(4)}(p,s,p',s';p+l,s,p'-l,s') &= \langle \psi_s(p)\psi_{s'}(p')\bar{\psi}_{s'}(p+l)\psi_s(p'-l)\rangle \\ &= \frac{\sum_n \frac{1}{n!}\langle \psi_s(p)\psi_{s'}(p')\bar{\psi}_{s'}(p+l)\psi_s(p'-l)\,(-S_I)^n\rangle_0}{\sum_n \frac{1}{n!}\langle (-S_I)^n\rangle_0}\,. \end{aligned}$$

Zeroth order: The zeroth-order term is simply

$$\begin{aligned} G^{(4)}(p,s,p',s';p+l,s,p'-l,s') &= \langle \psi_s(p)\psi_{s'}(p')\bar{\psi}_{s'}(p+l)\psi_s(p'-l)\rangle_0 \\ &= G_0^{(2),s}(p)G_0^{(2),s'}(p')\delta_{l,0} \\ &\quad -\delta_{ss'}G_0^{(2),s}(p)G_0^{(2),s'}(p')\delta_{p,p'-l} \end{aligned} \tag{15}$$

Diagrammatically, these two terms are disconnected Green's functions and hence do not contribute to $G_c^{(4)}(p,\uparrow,p'\downarrow;p+l,\uparrow;p-l,\downarrow)$ or the corresponding vertex.

First order: If we go to first order, for the connected contributions to be nonzero in the simplified case of our Hubbard interaction, we should choose $s =\uparrow$ and $s' =\downarrow$ or vice versa. Then we get

$$G_c^{(4)}(p,\uparrow,p'\downarrow;p+l,\uparrow;p-l,\downarrow) = -\frac{UT}{N} G_0^{(2)}(p) G_0^{(2)}(p') G_0^{(2)}(p+l) G_0^{(2)}(p'-l)\,.$$

So, comparing with Eq. 12 we see that the first order vertex is just $-UT/N$. The 'trivial' factor T/N is simply due to the representation in Matsubara frequency and momentum space.

Second order: Here the contributions are in general more involved and contain several terms that we will identify later as five possible one-loop diagrams. Here, for the restricted spin structure of the onsite interaction, some particle-hole-terms cancel and we obtain after using the wavevector and frequency conservation

$$\begin{aligned} O(U^2) &= G_0(p)G_0(p')G_0(p+l)G_0(p'-l)\cdot \\ &\quad \cdot\frac{T}{N}U^2\left[T\sum_k G_0(k)G_0(-k+p+p') + T\sum_k G_0(k)G_0(k+p'-l+p)\right] \\ &\quad + \text{corrections on external legs}\,. \end{aligned} \tag{16}$$

The first term in the brackets comes from a so-called particle-particle diagram, where both lines in between the two interactions go in the same direction, the second term is a (crossed) particle-particle diagram, with opposite direction of the internal lines. For a general interaction with spin summation, there would be three additional particle-hole diagrams. Below we compute these diagrams in more detail, and find out that the can give rise to divergent corrections at low T.

In addition to these 2nd-order terms there are (one-particle-reducible) corrections on the external legs, where on order of U creates a self-energy insertion on one of the external Green's functions. All such contributions can then be summed up in higher orders to give the full Green's functions on the external legs, as in Eq. 12. Here we focus on the one-particle-irreducible corrections, i.e. the changes from the 'bare' or first-order interaction vertex $-TU/N$ to the full interacting vertex due to multiple scattering processes.

In the next order, U^3, we either get products of one-loop diagrams, or true two-loop diagrams with two overlapping internal summations. The two-loop diagrams are usually less important in dimensions more than one, as they do not lead to divergent corrections at low T. The products of one-loop diagrams can be summed up to some extent, as will be done in single-channel summations in the following.

Another general observation that we can make is that the square bracket in Eq. 16 depends on the external leg variables. This means that through the perturbative corrections, the vertex will acquire a wavevector and frequency dependence, i.e. become nonlocal in space and imaginary time, even if the bare interaction is just a constant onsite interaction U.

Before continuing, let us study the one-loop particle-particle and particle-hole bubbles in more detail. For later use, and for getting clearer results, we will compute the diagrams with an infrared cutoff Λ imposed on the bare Green's function that cuts out the modes near the Fermi level for $|\epsilon(\vec{k})| \leq \Lambda$. The full results can be obtained letting $\Lambda \to 0$. The free propagator is then

$$G_{0,\Lambda}(i\omega,\vec{k}) = \frac{\Theta(|\epsilon(\vec{k})|-\Lambda)}{i\omega - \epsilon(\vec{k})}\,, \tag{17}$$

with the step function $\Theta(x) = 1$ for $x > 0$ and $= 0$ otherwise. The particle-hole diagram for incoming total frequency $i\nu$ (bosonic) and total wavevector $\vec{Q}$ is given by

$$L_\Lambda^{\rm PP}(i\nu, \vec{Q}) = \sum_{\vec{k}} T \sum_{i\omega_n} G_{0,\Lambda}(i\omega, \vec{k}) G_{0,\Lambda}(-i\omega + i\nu, -\vec{k} + \vec{Q}) \tag{18}$$

According to the standard rules of the Matsubara summation the particle-particle bubble then comes out as

$$L_\Lambda^{\rm PP}(i\nu, \vec{Q}) = \sum_{\vec{k}} \Theta(|\epsilon(\vec{k})| - \Lambda)\Theta(|\epsilon(-\vec{k} + \vec{Q})| - \Lambda) \, \frac{1 - n_F[\epsilon(\vec{k})] - n_F[\epsilon(-\vec{k} + \vec{Q})]}{-i\nu + \epsilon(\vec{k}) + \epsilon(-\vec{k} + \vec{Q})} \, . \tag{19}$$

Similarly we can derive for the particle-hole contributions

$$L_\Lambda^{\rm PH}(i\nu, \vec{Q}) = \sum_{\vec{k}} \Theta(|\epsilon(\vec{k})| - \Lambda)\Theta(|\epsilon(\vec{k} + \vec{Q})| - \Lambda) \, T \sum_{i\omega_n} G_{0,\Lambda}(i\omega, \vec{k}) G_{0,\Lambda}(i\omega + i\nu, \vec{k} + \vec{Q}) \tag{20}$$

after performing the Matsubara summation

$$L_\Lambda^{\rm PH}(i\nu, \vec{Q}) = \sum_{\vec{k}} \Theta(|\epsilon(\vec{k})| - \Lambda)\Theta(|\epsilon(\vec{k} + \vec{Q})| - \Lambda) \, \frac{n_F[\epsilon(\vec{k})] - n_F[\epsilon(\vec{k} + \vec{Q})]}{i\nu + \epsilon(\vec{k}) - \epsilon(\vec{k} + \vec{Q})} \, . \tag{21}$$

Let us now concentrate on vanishing external frequency, $i\nu = 0$, and for the particle-particle-diagram $\vec{Q} = 0$. Then we get, using inversion symmetry $\epsilon(\vec{k}) = \epsilon(-\vec{k})$, which is a property of most dispersions,

$$\begin{aligned} L_\Lambda^{\rm PP}(0,0) &= \sum_{\vec{k}} \Theta(|\epsilon(\vec{k})| - \Lambda)\Theta(|\epsilon(\vec{k})| - \Lambda) \, \frac{1 - 2n_F[\epsilon(\vec{k})]}{2\epsilon(\vec{k})} \\ &= \int_{-W}^{-\Lambda} d\epsilon \, \rho(\epsilon) \frac{1 - 2n_F[\epsilon]}{2\epsilon} + \int_{\Lambda}^{W} d\epsilon \, \rho(\epsilon) \frac{1 - 2n_F[\epsilon]}{2\epsilon} \, . \end{aligned} \tag{22}$$

In the last step we have changed from the wavevector summation to an integral over the band energy and used that the cutoff functions cut out the region around the Fermi level with $|\epsilon| < \Lambda$. $\rho(\epsilon)$ is the density of states at energy ϵ. If the temperature is much smaller than Λ, we can neglect the thermal smearing of the Fermi function and combine the two integrals to

$$\begin{aligned} L_\Lambda^{\rm PP}(0,0) &= \int_{\Lambda}^{W} d\epsilon \, \rho(\epsilon) \frac{1}{\epsilon} \\ &\approx \rho(0) \, \ln \frac{W}{\Lambda} \, . \end{aligned} \tag{23}$$

$\rho(0)$ is the density of states at the Fermi level, and we have assumed that the variation of $\rho(\epsilon)$ can be neglected in the respective energy window.
For $T \ll W$ we can also obtain a simple estimate by approximating the Fermi function by 0 or 1 for $\epsilon > T$ or $\epsilon < -T$ respectively, and by $n_F(\epsilon) \approx \frac{1}{2}(1 - \epsilon/T)$ in between. The interval betwen $-T$ and T then gives a finite contribution for $T \to 0$, while the outer regions give together

$$\begin{aligned} L_{\Lambda=0}^{\rm PP}(0,0) &= \int_{T}^{W} d\epsilon \, \rho(\epsilon) \frac{1}{\epsilon} \\ &\approx \rho(0) \, \ln \frac{W}{T} \, . \end{aligned} \tag{24}$$

We learn that the particle-particle diagram is divergent at low T and small Λ. We face a **logarithmic divergence** which is quite common in many-body problems. This divergence at zero total incoming frequency and wavevector is common to all many-fermion system obeying $\epsilon(\vec{k}) = \epsilon(-\vec{k})$ in any spatial dimension as long as the density of states at the Fermi level is nonzero. The analysis of the particle-particle channel at zero total incoming frequency and wavevector was the 'worst-case' analysis, other contributions of this diagram type are usually smaller and regular in the low-scale or low-T limit.
Now let us ask if we can get similarly divergent terms in the particle-hole channel. In order to get a analogous expression we now need to have $\epsilon(\vec{k}+\vec{Q}) = -\epsilon(\vec{k})$ for some fixed $\vec{Q}$ on a dense set of $\vec{k}$ values, for $\vec{k}$ towards the Fermi level. This is usually referred to as **nesting**, as a whole wavevector region matches on another one with the same (absolute) value of the dispersion. Nesting of the type is rather easily found in one dimension with a linear dispersion near the Fermi level and $Q = \pm 2k_F$. In two dimensions, the 'Hubbard' hopping dispersion on the square lattice with $t' = 0$ and $\mu = 0$ also fulfills $\epsilon(\vec{k}+\vec{Q}) = -\epsilon(\vec{k})$ with the nesting vector $\vec{Q} = (\pi, \pi)$. In realistic situations, neither t' nor μ will be exactly zero together, but of course the behavior for this special parameter point has a finite extent such that the situation found here matters for a certain range of parameters.
For the perfectly nested situation with $\epsilon(\vec{k}+\vec{Q}) = -\epsilon(\vec{k})$, we get also $n_F[\epsilon(\vec{k}+\vec{Q})] = 1 - n_F[\epsilon(\vec{k})]$. Then the particle-hole diagram for $i\nu = 0$ becomes

$$\begin{aligned} L_\Lambda^{\rm PH}(\vec{Q}) &= \sum_{\vec{k}} \Theta(|\epsilon(\vec{k})| - \Lambda)\Theta(|-\epsilon(\vec{k})| - \Lambda)\, \frac{-1 + 2n_F[\epsilon(\vec{k})]}{2\epsilon(\vec{k})} \\ &\approx -\rho(0)\, \ln \frac{W}{\Lambda}\,. \end{aligned} \tag{25}$$

Hence, as can be seen already from the formula in the first line, the particle-hole diagram for this situation has the opposite value as the particle-particle diagram and also diverges at low T and Λ.
Let us now look at the particle-particle diagram for $\vec{Q} \to 0$ and $\nu = 0$. We get

$$\begin{aligned} L_{\Lambda=0}^{\rm PH}(0, \vec{Q}\to 0) &= \sum_k \frac{n_F[\epsilon(\vec{k})] - n_F[\epsilon(\vec{k}+\vec{Q})]}{\epsilon(\vec{k}) - \epsilon(\vec{k}+\vec{Q})} \\ &\to \sum_{\vec{k}} \left.\frac{\partial n_F[\epsilon(\vec{k})]}{\partial \epsilon}\right|_{\vec{k}} \qquad (26) \\ &\to -\rho_0 \quad \text{for} \quad T \to 0\,, \qquad (27) \end{aligned}$$

as the derivative of the Fermi function becomes a negative δ-function at the Fermi level in the $T = 0$ limit, and ρ_0 denoting the density of states at the Fermi level *per spin* in this case. At finite T, the δ-function is temperature-smeared, and the density of states is averaged over a region of width $\sim T$. Note that we we include an infrared cutoff $\Lambda \gg T$, the loop contribution vanishes. In the t-t' Hubbard model, the dispersion exhibits van-Hove singularities at the points $(\pi, 0)$ and $(0, \pi)$ where the density of states diverges logarithmically $\sim \log \frac{\epsilon - \epsilon_{\rm vH}}{\epsilon_0}$, with $\epsilon_{\rm vH} = \epsilon(0, \pi)$ and some constant ϵ_0. In the case where the van Hove energy $\epsilon_{\rm vH}$ is zero, also ρ_0 is divergent. This means that besides the singularity in the $\vec{Q} = (\pi, \pi)$ particle-hole channel, there is the possibility of singular behavior in the small-$\vec{q}$-particle-hole channel. The particle-hole bubble at (π, π) and the the particle-particle bubble at $(0, 0)$ even become $\log^2$-divergent in this case.

2.3 Higher orders

In general, it is not possible to give a closed expression for the full perturbation series to all orders in U with all diagrams included. It is however not too difficult to see that the higher order terms, say to order U^n will among others contain diagrams that are chains or 'ladders' of n particle-particle or particle-hole diagrams. The internal summations within each segment of this chain are independent and the value of the diagram is basically the nth power of the particle-particle bubble. Picking out only these contributions can be justified in special cases. For example, in more than one dimensions when the particle-particle diagram at $\vec{Q} = 0$ and $i\nu = 0$ is the only log-divergent term, we find that at any given order, the ladder or bubble-chain contributions will dominate, as other insertions take away at least one log-divergent factor.
More precisely, summing up the particle-particle ladder, we get

$$\begin{aligned}
&\Gamma^{(4)}_{\mathrm{PP-ladder}}(p,\uparrow,p'\downarrow;p+l,\uparrow,p-l,\downarrow)/(-T/N) \\
&= U - U^2 L^{\mathrm{PP}}_\Lambda(p+p') + U^3\left[L^{\mathrm{PP}}_\Lambda(p+p')\right]^2 + \dots \\
&= U\sum_n \left[-UL^{\mathrm{PP}}_\Lambda(p+p')\right]^n \\
&= \frac{U}{1+UL^{\mathrm{PP}}_\Lambda(p+p')}\,.
\end{aligned} \tag{28}$$

The result now strongly depends on the sign of U, in particular for $p+p'=0$ where the PP bubble diverges logarithmically. Here, $L^{\mathrm{PP}}_\Lambda(p+p'=0) \approx \rho_0 \ln \frac{W}{\max(\Lambda,T)}$. If $U>0$ the vertex becomes very small near this particular wavevector and frequency combination in the limit $T,\,\Lambda \to 0$. Then the log-divergence of the bubble diagram is no longer significant, it basically switches off a part of the interactions. However, if $U<0$, the ladder sum diverges at nonzero T_c with a power-law form $\sim |T-T_c|^{-1}$, and the effective interaction including multiple scattering becomes infinitely strong. This leads to the formation of bound states of total wavevector zero, also known as Cooper pairs. This phenomenon is exactly the Cooper instability that expresses the instability of the Fermi sea with respect to attractive interactions.
Similar to the Cooper channel, also the previously considered PH-contribution in the perturbation series for the vertex is produced in arbitrary powers with alternating sign, and this particular part of the series can be summed to give (using $q = p'-l-p$)

$$\begin{aligned}
\Gamma^{(4)}_{\mathrm{PH-ladder}}(p,\uparrow,p'\downarrow;p+l,\uparrow;p-l,\downarrow)/(-T) &= U - U^2 L^{\mathrm{PH}}_\Lambda(l) + U^3\left[L^{\mathrm{PH}}_\Lambda(q)\right]^2 + \dots \\
&= U\sum_n \left[-UL^{\mathrm{PH}}_\Lambda(q)\right]^n \\
&= \frac{U}{1+UL^{\mathrm{PH}}_\Lambda(q)}\,.
\end{aligned} \tag{29}$$

In our language, $L^{\mathrm{PH}}_\Lambda(q)$ is negative. This means that the sign considerations from the Cooper case are exactly reversed. Now a positive U leads to an enhancement of the vertex, and if

$$UL^{\mathrm{PH}}_\Lambda(q) < -1 \tag{30}$$

we get an instability. As standard susceptibilities (i.e. expectation vales of two fermion bilinears) are contained in the two-particle Green's function, a divergence of the vertex $\Gamma^{(4)}$ corresponds to some divergent susceptibility. Here, the crossed particle-hole ladder for the given

spin combination appears in the magnetic susceptibility, and a divergence a wavevector transfer $\vec{q}$ can be understood as spin-density-wave instability. The criterion in Eq. 30 is a generalized Stoner criterion for magnetic ordering.
For generic dispersions, the particle-hole bubble remains finite, and a critical interaction strength $U_c \geq 1/|L_\Lambda^{\rm PH}(q)|$ is needed for an instability. Yet, as argued, for particular band dispersions, also the particle-hole diagram can diverge logarithmically (or even stronger) at zero frequency transfer and certain wavevector transfers.
Hence we see that separate 'single-channel' summations for the interaction verte can lead to different instabilities. Of course the full perturbation series contains in addition many other diagrams, and in particular diagrams, where possibly divergent PP and PH bubbles, e.g. those at $\vec{q} = \vec{Q}$ and those at $\vec{q} \approx 0$, mix. Then the question is whether these contributions will significantly alter the expectations from the single-channel summations. Furthermore, one would like to know which instability is actually the strongest or leading and has the highest critical temperature.
For this purpose one should sum all diagrams, or at least all one-loop diagrams and products of them together. The latter scheme goes under the name 'parquet summation' and has been performed for special cases (like in $D = 1$). In general however, the renormalization group framework is considered a more viable tool to accomplish this joint and unbiased summation of all possibly dangerous diagrams.

3 Functional RG formalism for lattice systems

3.1 Flow equation for the interaction vertex

The functional renormalization group method has been introduced in this spring school e.g. in the article by V. Meden. We directly jump to Equation (65) in his article, the flow equation for the two-particle vertex. In our notation it becomes

$$\begin{aligned}
\frac{d}{d\Lambda}\Gamma_\Lambda^{(4)}(K_1', K_2'; K_1, K_2) &= \sum_{P_1,P_1'}\sum_{P_2,P_2'} G^\Lambda(P_1,P_1')\, S^\Lambda(P_2,P_2') \\
&\times \Big\{\Gamma_\Lambda^{(4)}(K_1', K_2'; P_1, P_2)\Gamma_\Lambda^{(4)}(P_1', P_2'; K_1, K_2) \\
&- \Big[\Gamma_\Lambda^{(4)}(K_1', P_2'; K_1, P_1)\Gamma_\Lambda^{(4)}(P_1', K_2'; P_2, K_2) + (P_1 \leftrightarrow P_2, P_1' \leftrightarrow P_2')\Big] \\
&+ \Big[\Gamma_\Lambda^{(4)}(K_2', P_2'; K_1, P_1)\Gamma_\Lambda^{(4)}(P_1', K_1'; P_2, K_2) + (P_1 \leftrightarrow P_2, P_1' \leftrightarrow P_2')\Big]\Big\} \\
&- \sum_{P,P'} S^\Lambda(P,P')\,\Gamma_\Lambda^{(6)}(K_1', K_2', P'; K_1, K_2, P)\ . \qquad (31)
\end{aligned}$$

Here, the multi-index K comprises again $k = (i\omega_n, \vec{k})$ and the spin quantum number s. The two major approximations that are done in most fRG applications to interacting fermions on 2D lattices are now[5, 12]:
a) Neglecting the flow of the self-energy. The self-energy has a similar fRG flow equation that involves the two-particle vertex. While this equation can be integrated and hence the self-energy can be studied, feeding the self-energy back into the flow equation for the two-particle vertex is technically challenging, and has only been done partially in small number of works (see Refs. [13] or [14], in more extent it is possible in the spin-fRG[9]). Furthermore, the general

experience is that in most situations the flow of the self-energy does not alter the type of the leading instability. So, for establishing the main ordering tendencies it can be neglected.
b) Neglecting the the three-particle vertex $\Gamma^{(6)}_\Lambda(K'_1, K'_2, P'; K_1, K_2, P)$. This approximation is again mainly dictated by feasibility. The three-particle vertex depends on 5 momenta and 5 frequencies and is hence a rather involved object. It may be important in describing true strong-coupling physics like Mott-insulating states, but it is known that conventional long-range order can be described (at least qualitatively) without taking $\Gamma^{(6)}_\Lambda(K'_1, K'_2, P'; K_1, K_2, P)$ into account. Some effects of the six-point vertex have been considered in multiband situations[15].
Note that in Eq. 31 without the three-particle vertex there are several distinct contributions involving two two-particle vertices, corresponding to the familiar particle-particle, direct particle-hole, and crossed particle-hole channel, respectively, as shown diagrammatically in Fig. 1. Therefore, in the fRG all possibly important diagrams are treated together, and the competition and interplay between the different ordering tendencies is captured with any bias.

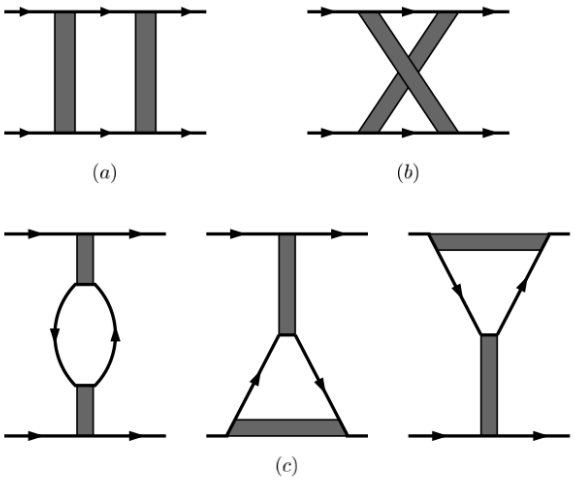


Fig. 1: *Contributions to the flow of the two-particle vertex with particle-particle (diagram a)) and particle-hole channel b) for the crossed channel, and the 'direct' particle-hole diagrams in c).*

For our translationally invariant systems, the Green's function in these terms is diagonal in the multi-index,

$$G^\Lambda_0(K, K') = \delta_{KK'}\, G^\Lambda_0(i\omega, \vec{k}) = \delta_{KK'}\, \frac{\Theta(|\epsilon(\vec{k})| - \Lambda)}{i\omega - \epsilon(\vec{k})}\,, \tag{32}$$

with the step function $\Theta(x) = 1$ for $x > 0$ and $= 0$ otherwise. This means that the low-energy modes below λ are set to zero. The fRG flow describes the change of the two-particle vertex when Λ is lowered. The subscript $_0$ signals that the self energy has been dropped. S^Λ is a single-scale propagator, which is with our approximations simply

$$S^\Lambda(K, K') = \delta_{KK'}\, S^\Lambda(i\omega, \vec{k}) = \delta_{KK'}\, \frac{\delta(|\epsilon(\vec{k})| - \Lambda)}{i\omega - \epsilon(\vec{k})}\,, \tag{33}$$

If the self-energy is indeed included, it has been shown to be advantageous to use the so-called Katanin replacement[16, 5].

The general interaction vertex $\Gamma^{(4)}_\Lambda$ is fully antisymmetric in the incoming legs K_1, K_2 and the outgoing legs K_1', K_2', hence we have for example

$$\Gamma^{(4)}_\Lambda(K_1,K_2;K_1',K_2') = -\Gamma^{(4)}_\Lambda(K_2,K_1;K_1',K_2') = \Gamma^{(4)}_\Lambda(K_2,K_1;K_2',K_1')\,. \tag{34}$$

Moreover, if spin is a conserved quantity as in the models considered here, the spin quantum numbers of in- and outgoing particles, s_1, s_2 and s_1', s_2' should add up to the same. This means that the vertex can have to the two possible contributions

$$\Gamma^{(4)}_\Lambda(K_1,K_2;K_1',K_2') = V_\Lambda(k_1,k_2;k_1',k_2')\delta_{s_1,s_1'}\delta_{s_2,s_2'} + \tilde{V}_\Lambda(k_1,k_2;k_1',k_2')\delta_{s_2,s_1'}\delta_{s_1,s_2'} \tag{35}$$

Here we have used that the multi-index K comprises $k = (i\omega_n, \vec{k})$ and the spin quantum number s. The antisymmetry under exchange of K_1 and K_2 gives now

$$V_\Lambda(k_1,k_2;k_1',k_2') = -\tilde{V}_\Lambda(k_2,k_1;k_1',k_2')\,. \tag{36}$$

Hence the vertex can be written as[12]

$$\Gamma^{(4)}_\Lambda(K_1,K_2;K_1',K_2') = V_\Lambda(k_1,k_2;k_1',k_2')\delta_{s_1,s_1'}\delta_{s_2,s_2'} - V_\Lambda(k_2,k_1;k_1',k_2')\delta_{s_2,s_1'}\delta_{s_1,s_2'}\,. \tag{37}$$

In particular, is is enough to compute the vertex for the case $s_1 = s_1' \neq s_2 = s_2'$, where only the first term contributes and equals the so-called coupling function $V_\Lambda(k_1,k_2;k_1',k_2') = V(k_1,k_2,k_1')$. The coupling function has the same spin for the first incoming and the first outgoing particle. It can be represented diagrammatically as a rectangle, where spin in conserved along the short edge. This notation is used in in Fig. 1.
The expression (37) for the vertex in terms of the coupling function can be inserted into the flow equations. This then gives the 5 diagrams shown in Fig. 1. With these, the flow equation for $V_\Lambda(k_1,k_2,k_3)$ with $k_4 = k_1 + k_2 - k_3$ reads[12]

$$\frac{d}{d\Lambda}V_\Lambda(k_1,k_2,k_3) = \mathcal{T}_{PP,\Lambda} + \mathcal{T}^d_{PH,\Lambda} + \mathcal{T}^{cr}_{PH,\Lambda} \tag{38}$$

with the one-loop particle-particle contributions $\mathcal{T}_{PP,\Lambda}$ and the two different particle-hole channels $\mathcal{T}^d_{PH,\Lambda}$ and $\mathcal{T}^{cr}_{PH,\Lambda}$ where

$$\begin{aligned}
&\mathcal{T}_{PP,\Lambda}(k_1,k_2;k_3,k_4) = \\
&\quad\int dk\, V_\Lambda(k_1,k_2,k)\, L(k,-k+k_1+k_2)\, V_\Lambda(k,-k+k_1+k_2,k_3)
\end{aligned} \tag{39}$$

$$\begin{aligned}
&\mathcal{T}^d_{PH,\Lambda}(k_1,k_2;k_3,k_4) = \\
&\quad\int dk \Big[-2V_\Lambda(k_1,k,k_3)\, L(k,k+k_1-k_3)\, V_\Lambda(k+k_1-k_3,k_2,k) \\
&\qquad +V_\Lambda(k_1,k,k+k_1-k_3)\, L(k,k+k_1-k_3)\, V_\Lambda(k+k_1-k_3,k_2,k) \\
&\qquad +V_\Lambda(k_1,k,k_3)\, L(k,k+k_1-k_3)\, V_\Lambda(k_2,k+k_1-k_3,k)\Big]
\end{aligned} \tag{40}$$

$$\begin{aligned}
&\mathcal{T}^{cr}_{PH,\Lambda}(k_1,k_2;k_3,k_4) = \\
&\quad\int dk\, V_\Lambda(k_1,k+k_2-k_3,k)\, L(k,k+k_2-k_3)\, V_\Lambda(k,k_2,k_3)
\end{aligned} \tag{41}$$

In these equations, the product of the two internal lines in the one-loop diagrams is $L(k,k') = S_\Lambda(k)G^\Lambda(k') + G^\Lambda(k)S_\Lambda(k')$.

3.2 Choice of flow parameter

Although we have made reference to a momentum-shell cutoff in the discussion of the 1PI fRG equations until now, as in the article by V. Meden, it is not difficult to see that the same equations can be written down for any other continuously variable parameter in the quadratic part of the action, provided that this parameter does not appear in the interaction term. This indicates how flexible and powerful these equations are. Here we want to discuss briefly which different flow parameters have been used previously to explore the instabilities of Hubbard-type lattice system. We mention however that in other contexts the freedom to choose appropriate flow parameters has been exploited as well (see e.g. Refs. [17, 18]).
The guiding principle for a good flow parameter in our context is that it should enable us *a)* to approach a specific infrared singularity in a controlled way (regulator property) and *b)* to include all other (possibly also singular) tendencies during the flow (unbiasedness). Then the fRG will give a more realistic picture than perturbation expansions that single out one dominant channel. As described before, in diagrammatic bubble or ladder summations, the singularities normally arise due to a pile-up of logarithms. These are roughly given by $g\log[W/\max(T,\Lambda)]$ (after Matsubara frequency summation) with a coupling constant g, bandwidth W and lower energy cutoff Λ. Hence, if we want to build in such dangerous terms step by step, we can either vary Λ, the temperature T, or the coupling g. Cutting out low Matsubara frequencies should have a similar effect as increasing T. This gives the following options.

- **Momentum-shell RG:** One writes for the free propagator

$$G_0^{\Lambda}(i\omega,\vec{k}) = \frac{C(|\epsilon(\vec{k})|-\Lambda)}{i\omega-\epsilon(\vec{k})}\,, \tag{42}$$

  with a cutoff function $C_\Lambda[\epsilon(\vec{k})]$, which can be chosen conveniently as a sharp step function for analytical manipulations (as above) or as a smoothened step function for numerical treatments. Then the Green's function is suppressed for modes with $|\epsilon(\vec{k})| \leq \Lambda$. The same holds for the single-scale propagator, hence only high-energy modes show up in the loop diagrams. The momentum-shell RG is the widely-used standard.

  A serious drawback of the momentum-shell cutoff RG is the non-uniformity in the RG scale at which one-loop particle-hole (PH) processes with different wavevector transfers $\vec{q}$ are included at low T. PH processes with small $\vec{q}$, i.e. a particle at wavevector $\vec{k}$ and a hole at $\vec{k}+\vec{q}$, can only occur in the vicinity of the Fermi surface. For $T\to 0$ the support of the PH bubble for $\omega=0, \vec{q}\to 0$ shrinks to a temperature-smeared δ-function on the Fermi surface with width $\sim T$. Thus, the $\vec{q}\to 0$ PH-modes are integrated out only for cutoff $\Lambda \leq T$, even if the density of states is divergent, and these processes give a singular contribution for $T\to 0$. However, in a coupled flow with various other tendencies, the flow normally diverges *before* we get down to $\Lambda \sim T$. The flow has to be stopped and the $\vec{q}\to 0$ PH pairs did not have the chance to contribute to the flow *by construction*. Therefore, the cutoff-flow is still biased to some extent and approaches are needed to study the influence of the $\vec{q}\to 0$ particle-hole excitations.

  The following three schemes work without a sharp cutoff and allow for a uniform inclusion of the important one-loop processes.

- **Temperature flow:** The temperature-flow scheme[19] uses the temperature T appearing in the quadratic part of the action as flow parameter. Before the T-derivative can be taken,

the fermion fields need to be rescaled in order to remove the T-dependence appearing in front of the interaction term in the original action. Then the theory is defined at some high $T \sim T_0$ in terms of its two-point and four-point vertex. Here, we simply choose the free propagator and the local Hubbard repulsion as the initial values at T_0. This should be reasonable for sufficiently high $T_0 \sim$ bandwidth, as all perturbative corrections decay with a negative power of T. In the approximate version without $\gamma_T^{(6)}$ and self-energy corrections, the right hand side of the flow equation for the interaction is just given by the T-derivatives of the one-loop diagrams. This flow scheme basically reproduced the antiferromagnetic and d-wave pairing instabilities found in the t-t'-Hubbard model using the momentum-shell schemes, but also added a ferromagnetic regime near van Hove filling and triplet pairing instabilities for larger values of $|t'|$.

- **Interaction flow:** In the interaction flow scheme[20] we first multiply Q with a scale factor $1/g$ and split this in two, yielding

$$Q_g = T \sum_{i\omega_n, \vec{k}} \bar{\psi}_k g^{-1/2} \left[-i\omega_n + \epsilon(\vec{k}) \right] \psi_k g^{-1/2} \,. \tag{43}$$

  g will be the flow parameter. We can absorb the factor $1/g$ in rescaled fields $\tilde{\psi}, \bar{\tilde{\psi}}$ defined as $\tilde{\psi} = g^{-1/2}\psi$. With this the interaction term gets an extra factor g^2 when written in terms of the new fields:

$$V_g^{(4)} = \frac{1}{2N} \sum_{\substack{k,k',k+q \\ s,s'}} g^2 V(k,k',k+q) \bar{\tilde{\psi}}_{k+q,s} \bar{\tilde{\psi}}_{k'-q,s'} \tilde{\psi}_{k',s'} \tilde{\psi}_{k,s} \,. \tag{44}$$

  We observe that changing the scale factor $1/g$ in Q_g corresponds to changing the strength of the bare interactions. The rescaled fermions $\tilde{\psi}$, $\bar{\tilde{\psi}}$ describe a system with a bare interaction strength $g^2\, V$. Now we can start at $g = 0^+$, i.e. at infinitesimally small bare interaction, and use the flow equations for the ψ-theory to integrate up to the desired bare interaction, reached at $g = 1$. We can also stop the flow at any other value of g, with the functions $g\Sigma$ and $g^2 V_g(k,k',k+q)$ being the self-energy and interacting vertex function for the bare interaction $g^2 V(k,k',k+q)$. One can call this scheme *interaction flow* (IF), or simply flat cutoff flow, as all modes are suppressed by the same factor g. Now singularities on the right hand side of the flow equation are not regularized by the flow parameter. Thus, the IF scheme has to be performed at $T > 0$, when the one-loop diagrams are bounded. Using this scheme, the ground state phase diagram of the 2D t-t' Hubbard model comes out basically the same as in the T-flow.

- **Smooth frequency cutoff:** Another regulator that does not ignore small-wavevector transfer processes in the particle-hole channel and is hence capable of describing ferromagnetic instabilities as well is the so-called Ω-scheme introduced by Husemann and Salmhofer[21]. Here one uses a smooth cutoff on the Matsubara frequency axis, given by $C_\Omega(\omega_n) = \omega_n^2/(\omega_n^2 + \Omega^2)$, and the flow goes from large Ω damping away all modes down to $\Omega = 0$ where the full free propagator is obtained. The benefit of the simple functional dependence on ω_n and Ω is that this cutoff function still permits one to determine the Matsubara sums in the loop diagrams analytically. Various works with the related singular-mode-fRG (SMFRG) use a sharp frequent cutoff[22].

4 Implementation of the fermionic fRG for two dimensional lattices

Besides that the fRG equations above are derived from exact flow equations for generating functionals, the name *functional renormalization group* reflects that at least for many-fermion systems with a Fermi surface, one can in general no longer reduce the parametrization of the interaction vertex to a handful of running coupling constants. Rather, the functional dependence of the interactions on the location on the Fermi surface or more generally in the Brillouin zone can turn out to be relevant, i.e. grow strongly under the RG flow. Hence keeping the functional dependence of the interaction vertex around the Fermi surface(s) is important. This leads to RG equations for functions of wavevectors, i.e. functional differential equations. The two-fermion interaction vertex, e.g., depends on three wavevectors, while the fourth is fixed by wavevector conservation on the lattice. Note that in principle also the frequency dependence of the interactions can be studied, but in the works described here the frequency dependence is mainly neglected. Here we describe two ways to deal with the wave vector dependence.

4.1 N-patch schemes

The wavevector dependence is typically treated by a so-called N-patch discretization, where the Brillouin zone is divided up into patches containing Fermi surface segments. This was first used in this context by Zanchi and Schulz[23]. Within these patches, the wavevector dependence is ignored again, and the value for the patch region near the Fermi surface is computed and used at all wavevectors inside the patch. However, as described below, a sufficient number of patches allows one to draw a number physically relevant conclusions. The setup of the patches around the Fermi surfaces is shown in Fig. 2. It is guided by the idea that one wants to obtain an effective interaction for quasiparticles near the Fermi surface, and the wavevector-dependence tangential to the Fermi surface is relevant in standard examples. The dependence normal to the Fermi surface does usually not alter the leading flow. Various other patching schemes, like square patchings of the Brillouin zone have been investigated as well, without unexpected or problematic differences in the results[20, 24].

Let us now focus on the mainly used approximation which consists (besides the truncation explained earlier) in neglecting the self-energy feedback. Then only the interaction vertex is considered. Further, we ignore the frequency dependence. By the N-patch construction, the functional RG equations for $V_\Lambda(\vec{k}_1, \vec{k}_2, \vec{k}_3)$ are converted into N^3 coupled nonlinear differential equations for the components of the discretized interaction vertex. As at least the particle-particle channel contains singular diagrams for $T \to 0$, the solution of these fRG equations without self-energy corrections typically results in a *flow to strong coupling* where one or more families of interaction processes run to large absolute values. This is accompanied by the build-up of sharp structures in wavevector-space, corresponding to the formation of longer-range interactions. Then the flow has to be stopped when the coupling strength has exceeded the scale of the band width, as then self-energy corrections and possibly also the truncation error would become important. From this scale where the run-away flow occurs one can derive an estimate for critical scales for long-range ordering or gap openings. Furthermore from the analysis which class of coupling function diverges most strongly, one can infer the dominant tendencies toward long-range ordering.

A tool for assessing physical information, e.g., on the gap function in the case of a Cooper

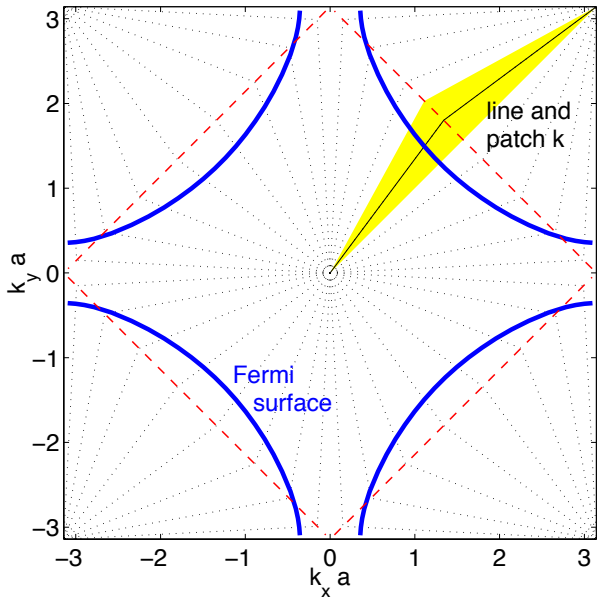


Fig. 2: *N-patch discretization of the Brillouin zone for the one-band Hubbard model on the 2D square lattice. The colored region is a patch in which the coupling function is approximated by a constant.*

instability is the diagonalization of the effective interaction in the Cooper (zero total incoming wavevector) channel, $V_\Lambda(\vec{k}, -\vec{k} \to \vec{k}', -\vec{k}')$ at a scale Λ near the instability[25, 6]. By studying the linearized gap function, one can see that the pairing gap function would adopt the $\vec{k}$-dependence of the eigenvector belonging to the most negative eigenvalue of $V_\Lambda(\vec{k}, -\vec{k} \to \vec{k}', -\vec{k}')$. This allows one to determine the structure of the pairing around the Fermi surface beyond simple symmetry statements. This reasoning can of course be extended to other order parameters.

4.2 Channel decomposition

As shown in Sec. 5 and also foreseeable from the second-order perturbation theory, the flow to strong coupling typically occurs for components of the coupling function $V_\Lambda(\vec{k}_1, \vec{k}_2, \vec{k}_3)$ with particular vales of the the total wavevector $\vec{k}_1 + \vec{k}_2$ or a particular wavevector transfer $\vec{k}_2 - \vec{k}_3$. While the dependence of the coupling function of these combinations of wavevectors is rather strong, there is usually a weaker dependence on the precise location of the external legs on the Fermi surface which can be well described with lower lattice harmonics such as a constant or a $d_{x^2-y^2}$-formfactor. This led Husemann and Salmhofer[21] to propose an alternative description of the coupling function,

$$V_\Lambda(\vec{k}_1, \vec{k}_2, \vec{k}_3) = U - \Phi^\Lambda_{\rm SC}(\vec{k}_1, \vec{k}_3, \vec{k}_1 + \vec{k}_2) + \Phi^\Lambda_{\rm M}(\vec{k}_1, \vec{k}_2, \vec{k}_3 - \vec{k}_1) \quad (45)$$

$$+\frac{1}{2}\Phi^\Lambda_{\rm M}(\vec{k}_1, \vec{k}_2, \vec{k}_2 - \vec{k}_3) - \frac{1}{2}\Phi^\Lambda_{\rm K}(\vec{k}_1, \vec{k}_2, \vec{k}_2 - \vec{k}_3) \quad (46)$$

where now the new functions $\Phi_{\rm SC}$, $\Phi_{\rm M}$ and $\Phi_{\rm K}$ are treated with a fine resolution for the last entry, and where formfactors are used to describe their dependence on the first two entries. Moreover, the particle-particle channel is selected to drive the flow of the pairing part $\Phi_{\rm SC}$, the

magnetic (crossed particle-hole) channel drives magnetic part $\Phi_{\rm M}$, and the another combination of particle-hole terms drives the charge part $\Phi_{\rm K}$. All this does not involve any additional approximations, but for larger parameter region it turned out[21] to be an efficient additional approximation to use only s-wave (constant) and $d_{x^2-y^2}$-formfactors to describe the dependence of the Φ-functions on their two first arguments. Therefore only one wavevector dependence, namely on the last argument, for the Φ-functions needs to be treated numerically. This reduces the computational effort and allows for a better description of the strong dependence on these wavevector combinations. The results for the repulsive Hubbard model at van Hove-filling on the 2D square lattice are briefly described in Sec. 5. We expect that this alternative formulation will prove useful also in the study of flows into the symmetry-broken state, and in order to capture frequency dependences better. Very similar decompositions for the frequency structure of the vertex function had already been used in the context of impurity models[26]. Related ideas have been implemented in pendently and successfully by Q.H. Wang and colleagues[22]. Another application of the idea without the use of explicit form factors has been proposed in Ref. [24].

5 Instabilities in two-dimensional lattice systems

Here we describe the results obtained by applying the N-patch fRG scheme to three interesting systems. The first system is the one-band Hubbard model on the 2D square lattice, whose phase diagram is of interest in connection with the high-T_c cuprates and other unconventional superconductors. As a second example we discuss fRG studies of multiband Hubbard-type systems on the same lattice with direct relevance to the newly discovered iron pnictide (and chalcogenide) superconductors. Finally we sketch recent results obtained for layered graphene.

5.1 Two-dimensional Hubbard model near half filling

The 2D Hubbard model is one of the most-studied theoretical models in context with high-temperature superconductivity in the layered copper-oxide materials. The Hamiltonian was given in Eq. 1.

While the original derivation of an effective one-band model for copper-oxide planes[27] implies a strong local Coulomb interaction U on the copper $d_{x^2-y^2}$-orbitals, the model has also been investigated for various reasons intensively at weaker couplings. Here we discuss the picture that is obtained by using the functional RG which is perturbative in U. With respect to the cuprates, studying the Hubbard model at weak to moderate U is relevant at least for cuprates in the overdoped regime. Here, experiments[28, 29] exhibit a rather well-defined 2D Fermi liquid with Fermi surfaces that can be parameterized by simple one-band tight-binding prescriptions.

Regarding the possible ground states of the one-band Hubbard model on the square lattice at weak to moderate U, the Fermi surface location and shape plays a dominant role. Let us now consider how the N-patch fRG-flows differ for different Fermi surfaces, parametrized by the second-nearest neighbor hopping t' and the chemical potential μ.

For $\mu = 0$ and $t' = 0$, one has a half-filled band with a perfectly nested Fermi surface. All electronic states that are occupied at $T = 0$, i.e. with $\epsilon(\vec{k}) < 0$, can be scattered by wavevector $\vec{Q} = (\pi, \pi)$ on unoccupied states, leading to a dominance of particle-hole-fluctuations with this nesting wavevector. An infinite-order summation of this particular particle-hole diagram would result in a divergent static spin susceptibility at $\vec{Q}$ for any $U > 0$ at sufficiently low tempera-

tures indicating the formation of an antiferromagnetic (AF) spin-density wave, as described in Subsec. 2.3. The basic fRG results for low T are shown in the left half of Fig. 3. In the upper plot the Fermi surface is shown together with $N = 32$ discretization points. In the middle plot we show the flow of two families of coupling constants. We can see that some of these lines flow to large values when the scale parameter Λ is lowered. This is an example for a *flow to strong coupling*. When the interaction reaches values larger than the band width the flow has to be stopped, as explained ind Sec. 4. Next to the divergence scale we can also read out from the data which classes of coupling constants grow most strongly. In this case, these are the dashed lines corresponding to interaction processes with wavevector transfer $\vec{Q} = (\pi, \pi)$ between $\vec{k}_2$ and $\vec{k}_3$. The maximal coupling constant that grows most strongly also belongs to this family. All members of this family flow to strong coupling in a rather similar way when Λ is reduced toward the critical scale $\Lambda_c \sim 0.16t$. Other families of couplings constants grow less strongly. Shown as well are as solid lines Cooper processes with zero total incoming wavevector. These processes show some growth as well, but lag behind the leading components. The lowest plot on the left shows a snapshot of the coupling constant when the first outgoing wavevector $\vec{k}_3$ is fixed at discretization point 1 near $(-\pi, 0)$ as function of the incoming wavevectors. We can clearly see two structures: one vertical line with strongly enhanced repulsive interactions at $k_2 = 24$ (corresponding to wavevector transfer $\vec{Q} = (\pi, \pi)$ between $\vec{k}_2$ and $\vec{k}_3$) with very little dependence on k_1, and another line at $k_1 = 24$ (corresponding to wavevector transfer $\vec{Q} = (\pi, \pi)$ between $\vec{k}_1$ and $\vec{k}_3$). These values show again only a weak dependence on k_2 and are roughly half as large as the vertical feature. Concentrating on these two features we arrive at the following effective interaction near the instability

$$H_\Lambda^{(\mathrm{SDW})} = \frac{V_{\mathrm{AF}}}{2} \sum_{\substack{\vec{k}_1,\vec{k}_2,\vec{k}_3 \\ s,s'}} \left[2\delta(\vec{k}_2 - \vec{k}_3 \pm \vec{Q}) + \delta(\vec{k}_1 - \vec{k}_3 \pm \vec{Q}) \right] c^\dagger_{\vec{k}_3,s} c^\dagger_{\vec{k}_4,s'} c_{\vec{k}_2,s'} c_{\vec{k}_1,s} \,,$$

where $\vec{k}_4$ is understood to given by $\vec{k}_1 + \vec{k}_2 - \vec{k}_3$ modulo reciprocal lattice vectors. A simple calculation shows that this interaction is equivalent to a long-ranged AF spin-spin interaction $H = -J \sum_{\langle i,j \rangle} e^{i\vec{Q}\cdot(\vec{R}_i - \vec{R}_j)} \vec{S}_i \cdot \vec{S}_j$ with spin operators given by the usual fermion bilinears and $J \propto V_{\mathrm{AF}}$. This effective Hamiltonian exhibits long-ranged AF order at sufficiently low T. In this sense, the fRG flow for these parameters clearly indicate the proximity to an AF ordered state or an AF spin-density wave (AF-SDW). Strictly speaking, this is the most we can infer at this stage. Whether long-range order actually occurs depends on the subleading terms and on the approximation errors collected on the way to this result. Note that the dynamics and interactions of emergent collective degrees of freedom are not appropriately captured in this approximation. Of course in two dimensions, a spontaneous breaking of a continuous symmetry should not occur at $T > 0$. On the other hand, most experimental systems described by our model would have additional small couplings in the third direction that change the situation in favor of our result. Moreover, as the leading instability is clearly exposed by this scheme, one could now resort to a bosonized description that allows one to treat the collective infrared physics much better (see e.g. Ref. [30] how the continuous symmetry breaking in two dimensions gets healed in the asymptotic flow).

Now let us consider a different Fermi surface with more curvature. For that we choose $t' = -0.3t$ and $\mu = -1.2$. Now the Fermi surface is curved but it still contains the van Hove points $(\pi, 0)$ and $(0, \pi)$. The resulting flow is illustrated in the right plots in Fig. 3. In the middle plot on the right we now see a dominance of the Cooper scattering processes with zero incoming total wavevector. Here some lines seem to diverge to large positive values, while

some other lines take very negative values. The sign structure is also visible quite nicely in the bottom figure on the right. Here the 1st outgoing wavevector $\vec{(k)}_3$ is again fixed at point 1 (near $(-\pi, 0)$) and the dependence on the incoming wavevectors around the Fermi surface is plotted. One observes diagonal lines of enhanced interactions, corresponding to zero total incoming wavevector $\vec{k}_1 + \vec{k}_2 = 0$. This pair scattering is attractive when the incoming pair $\vec{k}_1$, $\vec{k}_2$ near the same saddle point $(\pm\pi, 0)$ as the outgoing pair $\vec{k}_3, -\vec{k}_3$, and repulsive, when incoming and outgoing pairs are at different saddle points. This is exactly the formfactor $d(\vec{k}) = d_0(\cos k_x - \cos k_y)$ of a $d_{x^2-y^2}$-Cooper pairing instability, where the dominant interaction is given by what is obtained by only keeping these diagonal features in $V_\Lambda(\vec{k}_1, \vec{k}_2, \vec{k}_3)$,

$$H_\Lambda^{(d\mathrm{SC})} = V_{d\mathrm{SC}} \sum_{\vec{k},\vec{k}'} d(\vec{k}) d(\vec{k}') \, c^\dagger_{\vec{k}',\uparrow} c^\dagger_{-\vec{k}',\downarrow} c_{-\vec{k},\downarrow} c_{\vec{k},\uparrow} \, .$$

The spin-structure can be understood by forming the antisymmetric vertex again via Eq. 37. As the dominant features are symmetric with respect to interchange of k_1 and k_2, the equal-spin vertex vanishes. Again, this effective Hamiltonain can be solved by mean-field theory exactly. It has a d-wave paired ground state. This d-wave pairing instability was obtained by a number of research groups with different variations of fRG approaches (e.g. [23, 31, 32, 33, 19]). The parameter region for its occurrence is rather wide. It constitutes convincing evidence that the weakly coupled Hubbard model possesses a d-wave superconducting ground state. The pairing mechanism at these weaker couplings is well described as AF-spin-fluctuation induced. This is already visible in the bottom plot on the right hand side of Fig. 3. Here one can see that the d-wave pairing interaction on the diagonal lines with zero total incoming wavevector crosses a region with enhanced repulsive interactions, e.g. near $k_1 \sim 8, 9$ and $k_2 \sim 24, 25$. This enhancement is the broadened, due to Fermi surface curvature, version of the vertical SDW feature in the left lower plot for the fully nested case. We see that the enhancement to positive values fits perfectly into the sign structure of the d-wave. Studying the flow as a function of the cutoff Λ, one finds that the SDW features appear first and create the attractive component in the $d_{x^2-y^2}$-pairing channel. More drastically, one can put a filter into the fRG scheme that removes these SDW features by hand, e.g. by setting these interactions to their initial values. This way one also looses the d-wave pairing instability in the same way as the AF-SDW instability, showing the strong coupling between the channels. Another way to check this is to completely drop the particle-hole channels. In this case, no flow to strong coupling is found, as no attractive pairing components are generated, and as the SDW instability is excluded as well. If we only drop the particle-particle channel, we of course destroy the pairing instabilities, and the SDW instability takes a wider parameter space, at even higher critical scales comparable to those found in single-channel summations.

Close to the instability, the fRG data discussed so far showed rather clearcut single-channel instabilities, either in the AF-SDW or in the d-wave pairing channel, provided T is low enough. For higher $T > T_c$ the flow remains finite, and the system should be in a metallic state. The critical scale at $T = 0$ is usually a good estimate for the critical temperatures T_c above which the flow remains finite, up to factors ~ 1. We can ask how the flow changes when we move from one to the other by changing the parameters. In principle there are two possibilities: One is that the flow to strong coupling changes gradually and the critical scale remains nonzero, with the SDW component getting weaker and the pairing component getting stronger continuously. The second possibility would be a quantum critical point where the critical scale for the run-away flow scale is suppressed to zero. As shown in the upper plots of Fig. 4 the change from

AF-SDW is of the first type, while for larger t' one finds indications for quantum critical point between d-wave pairing and ferromagnetism, at least in the fermionic fRG flows, as illustrated in the lower plots of Fig. 4. The main factor that causes this difference between the phase changes is the overlap of the interaction processes between the two respective channels.

The AF-SDW instability and the $d_{x^2-y^2}$ both require repulsive interaction processes between the two saddle points region near $(\pi, 0)$ and $(0, \pi)$. Hence the growth of one channel also supports the other channel to a large extent. If the Fermi surface remains in the the vicinity of the van Hove points for a larger range of μ, there is a rather wide crossover region where both channels, d-wave pairing and AF-SDW, grow strongly in a rather similar fashion. This is the case in the so-called saddle-point regime in the upper right plot of Fig. 4. Here, the mutual reinforcement of different interaction processes with initial and final states near the saddle-point regions is reminiscent of the flows to strong coupling in the half-filled two-leg Hubbard model, where d-wave pairing and AF-SDW channel are again driven in parts by the same processes. In this one-dimensional system the ground state does not select one these channels and does not develop any quasi-long-range order, but becomes a short-range correlated spin liquid with a single-particle gap[34]. This resemblance motivated interpretations of the flows in two dimensions as indications for a partial truncation of the Fermi surface in the saddle point regions[32, 35, 36], similar to what is argued to occur in the underdoped high-T_c cuprates. A controlled determination of the resulting strong-coupling state near the saddle-point region for these parameter regions is however still missing.

If the Fermi surface is further away from the saddle point regions, the strongest interaction processes with wavevector transfer (π, π) occur near the Brillouin zone diagonal where the $d_{x^2-y^2}$-formfactor is small. Then d-wave pairing and AF-SDW channels are rather weakly coupled. Correspondingly the energy scales for the two instabilities can be quite different. This is seen in the left upper plot of Fig. 4 for band fillings larger than one. Here the high-energy-scale AF-SDW instability gets cut off at a certain critical chemical potential and is replaced by a low-energy-scale $d_{x^2-y^2}$ pairing instability. Without further analysis it is difficult to determine the precise nature of this transition. However, as the coupling between the channels is rather weak, we expect that a more accurate calculation would yield first order transition.

The situation is again different for the transition from the d-wave pairing to the ferromagnetic instability, as illustrated in the lower plots in Fig. 4. Here the critical scale for the flow to strong coupling is suppressed strongly in the transition region, and the fermionic fRG flows (left lower plot) even suggest a quantum critical point. While it is difficult to fully understand these coupled non-linear flows of N^3 components of the coupling function using a simple argument, there is now one significant difference compared to the continuous crossovers described above. Let us again consider the interaction process that is common to both, d-wave pairing and ferromagnetic channel. The latter channel is driven be scattering processes with wavevector transfer $\vec{q}$ (between $\vec{k}_2$ and $\vec{k}_3$) going to zero. A closer analysis shows that these processes should be repulsive in order to cause a ferromagnetic instability, like for the usual Stoner criterion. This however means that also the processes with $\vec{q} \to 0$ that have zero total momentum and enter the Cooper channel should be repulsive. A positive pair scattering with small wavevector transfer is however incompatible with singlet Cooper pairing that needs attractive scattering for these wavevector combinations. In this sense, the overlap between the two ordering channels can either support ferromagnetism or d-wave pairing, but not both tendencies together, and the two channels compete already in the symmetric phase. We note as well that the ferromagnetic instability is rather fragile and limited to the neighborhood of van Hove filling. Upon doping away one find a smooth crossover to an instability with low critical scale and dominant p-wave

pairing interactions[19]. The plot on the right hand side of Fig. 4 is also consistent with the channel-decompsed flows in Refs. [21] and [13]. Here the fermionic vertex was decomposed into different functions in charge, spin and pairing channel and with s- and d-wave formfactors that depend only on one specific wavevector transfer or the total incoming wavevector (as discussed in the end of Sec. 4).
Summarizing these results on the one-band Hubbard model we see that the fRG is capable of deriving tentative phase diagrams with detailed descriptions of the wavevector- (and in principle also frequency-)structure of the effective interactions. Next to establishing a broader regime with clearcut instabilities toward phases with unconventional Cooper pairing, the method also shows that the effective interactions at the borders of these regimes are rather complex. To understand the physical meaning of these flows with several strong channels in the effective interaction and to relate them to observable phenomena is an interesting challenge for the future research.
Before going on, we want to mention recent advances in pushing the flow into the symmetry-broken states[37]. Here the flow of an anomalous self-energy is included and all modes can be integrated out, with a gapped single-particle spectrum as end result of the flow. The first application to the repulsive Hubbard model has just been completed[38].

5.2 Iron pnictides

A new field where fRG has contributed considerably are the newly discovered iron pnictide superconductors[1]. Here the applicability of a perturbative technique like the one described here may be even better, as the pncitides are certainly not as strongly correlated as the high-T_c cuprates. This can already be inferred from the experimental phase diagram, where one only finds metallic antiferromagnetic phases (for some pnictide materials this is missing), but never Mott insulating antiferromagnetism. Further, theoretical works that try to estimate the value of the iron-d orbital onsite-interaction strengths put the materials into the range of weak to moderate correlations[40, 41]. Regarding their electronic structure, the pnictides are more complex than the cuprates. At least three of the five iron d-orbital have non-negligible weight near the Fermi level[42, 43], and that these d-orbitals hybridize strongly with the neighboring arsenic p-orbitals. Therefore, the multi-band character has to be kept even very close to the Fermi level. The Fermi surface is divided into two hole pockets, centered around the origin of the Brillouin zone at $\vec{k} = 0$, and two electron pockets around $\vec{k} = (\pi, 0)$ and $\vec{k} = (0, \pi)$ in the unfolded zone corresponding to the small unit cell with one iron atom (or $\vec{k} = (\pi, \pi)$ in the folded zone corresponding to the large unit cell with two iron atoms). As pointed out early[42, 44], there is a potential nesting of electron- and hole pockets which enhances particle-hole susceptibilities with the wavevector connecting these pockets. In addition, depending on the parameters and approximations[45], there can be a third hole pocket at (π, π) in the unfolded zone.
The first fRG work on the pnictides was performed by the Berkeley group[25] for a five-band model. These authors obtained a sign-changing s-wave pairing instability driven by AF fluctuations as the dominant pairing instability. Further they found strongly anisotropic gaps around the electron pockets, with possibility of node formation. The basic structure of the phase diagram with the sign-changing pairing gap between electron- and hole-pockets can be understood already from simplified few-patch RG approaches[46]. This would however predict isotropic gaps around these pockets[47]. To understand the gap anisotropy one has to take into account the multi-orbital nature of the electronic spectrum in the iron pnictides, as was done by the

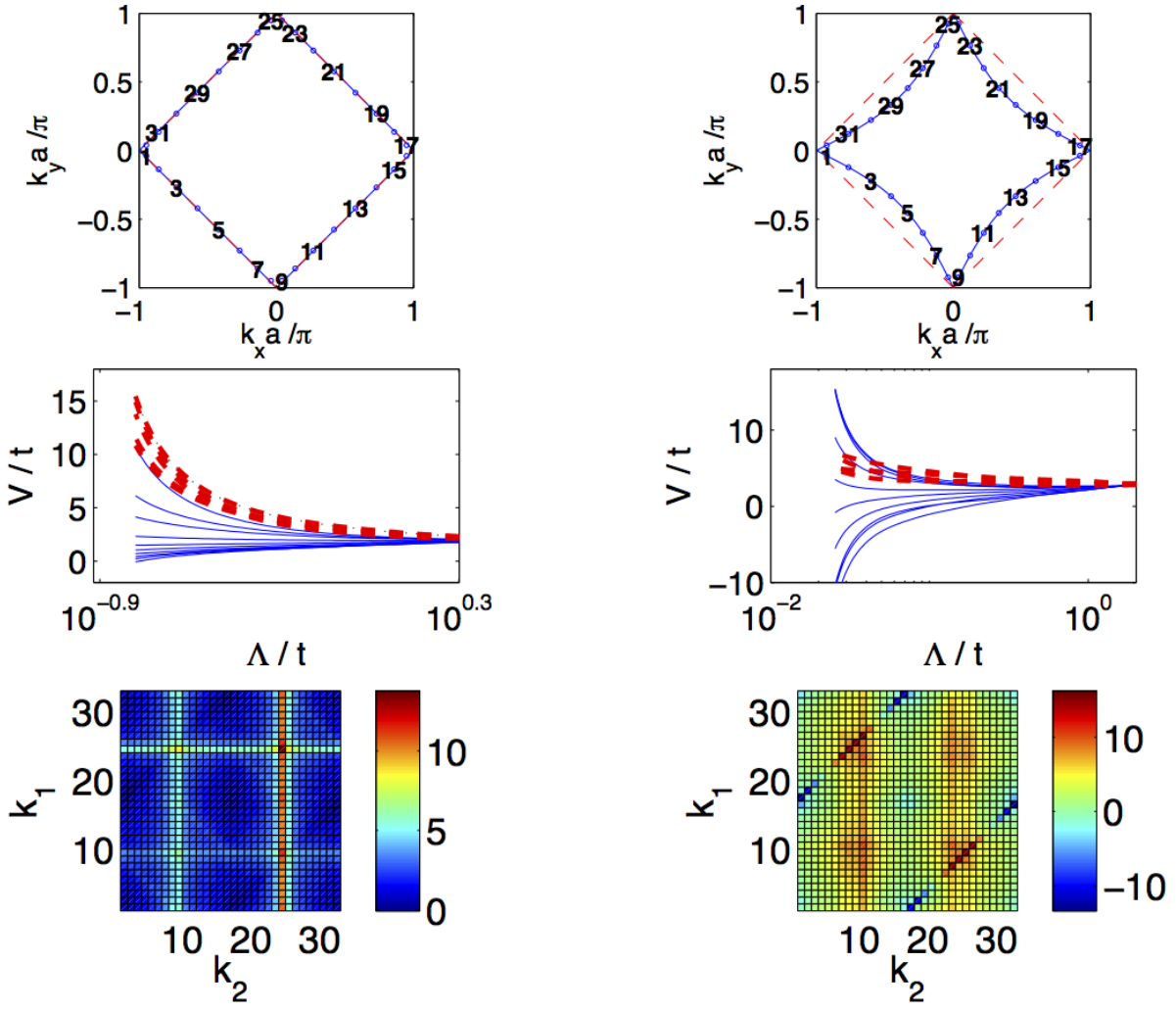


Fig. 3: *N-patch fRG data obtained with the momentum-shell 1PI fRG scheme for the repulsive Hubbard model on the 2D square lattice. Left plots: $\mu = 0$, $t' = 0$ and initial $U = 2t$, right plots: $\mu = 1.2t$, $t' = -0.3t$, $U = 3t$. Upper row: Fermi surfaces for the two cases and the $N = 32$ discretization points for the two incoming $\vec{k}_1$, $\vec{k}_2$ and the 1st outgoing wavevector $\vec{k}_3$. Middle row: Solid lines show the flows of components in the coupling function $V_\Lambda(k_1, k_2, k_3)$ corresponding to Cooper pair scattering with zero total incoming wavevector, $\vec{k}_1 + \vec{k}_2 = 0$, or $|k_1 - k_2| = N/2$ in terms of patch indices. The dashed lines correspond to processes in the AF-SDW channel with wavevector transfer $\vec{k}_2 - \vec{k}_3$ near $\pm(\pi, \pi)$. The flow is started at $\Lambda = 4t$ and goes to the left toward smaller Λ. Lower plots: Snapshots of the coupling function $V_\Lambda(k_1, k_2, k_3)$ near the instability with k_1 fixed at point 1 with k_1 and k_2 moving around the Fermi surface. The colorbars on the right indicate the values of the interactions in units of t.*

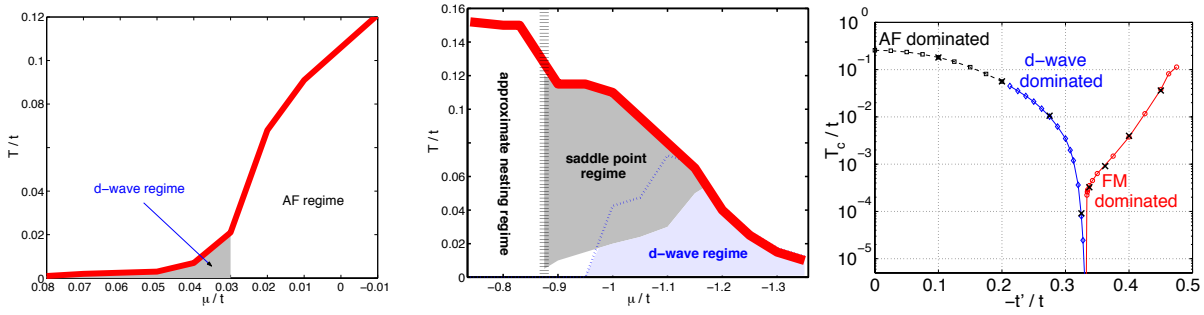


Fig. 4: *Leading instabilities as found by N-patch fRG in the t-t'-Hubbard model. Left plot: Critical scale T_c below which the fRG flow goes to strong coupling vs. chemical potential μ for band filling larger than unity, for $t' = -0.3t$ and $U = 3t$. There is a high-energy-scale AF SDW instability with a weaker $d_{x^2-y^2}$-wave pairing instability when the AF-SDW is cut off. Data from Ref. [39]. Middle plot: Data for the same t' and U on the 'hole-doped' side with band fillings smaller than one, from Ref. [32]. Now there is a broad crossover 'saddle point regime' between the nesting-driven AF-SDW instability and the $d_{x^2-y^2}$-wave pairing regime. Right plot: T_c vs. t' at the van Hove filling where the Fermi surface contains the points $(\pi, 0)$ and $(0, \pi)$. For large t' one finds a ferromagnetic instability. Data from Ref. [19] obtained with the T-flow scheme.*

Berkeley group in their initial study[25]. Let us start with a single-particle Hamiltonian in wavevector-orbital space

$$H = \sum_{\vec{k},s,o} h(\vec{k})_{oo'} c^\dagger_{\vec{k},o,s} c_{\vec{k},o',s} \tag{47}$$

where the matrices $h(\vec{k})_{oo'}$ take into account intra- and inter-orbital (density-density interactions and Hund's rule) terms for orbital index $o = o'$ or $o \neq o'$ respectively. s is the spin quantum number. The energy bands are obtained by a unitary transformation from orbital to band operators (index b), $c_{\vec{k},b,s} = \sum_o u_{bo}(\vec{k}) c_{\vec{k},o',s}$. For the various density-functional-theory-based tight-binding parameterizations of the band structure (e.g. [41, 48, 49]) of the d-dominated bands, three orbitals d_{xy}, d_{xz} and d_{yz} (in the coordinates of the small cell, with x and y-axis pointing towards the nearest iron neighbor) contribute significantly near the Fermi level. The simplest choice for the interaction between the electrons is to introduce orbital-dependent intra- and inter-orbital onsite repulsions, plus Hund's rule and pair hopping terms. While these local terms lead to wavevector-independent interactions in the orbital basis, parametrized by a tensor $V_{o1,o2,o3,o4}$, after the transformation to bands one arrives at a wavevector-dependent interaction function

$$\begin{aligned} V_{b1,b2,b3,b4}(\vec{k}_1, \vec{k}_2, \vec{k}_3, \vec{k}_4) &= \sum_{o1,o2,o3,o4} V_{o1,o2,o3,o4} \\ &\cdot \; u_{b1,o1}(\vec{k}_1) u_{b2,o2}(\vec{k}_2) u^*_{b3,o3}(\vec{k}_3) u^*_{b4,o4}(\vec{k}_4) \,. \end{aligned} \tag{48}$$

The combination of u_{bo}s behind the interaction tensor is sometimes called the 'orbital make-up'[48]. These prefactors cause in practice that already the initial interaction of the fRG flow exhibits a marked wavevector-structure which is then renormalized during the flow. It turns out that this orbital make-up has an essential influence on the competition between different channels in the flow and is responsible for the gap anisotropies found in the multi-band fRG studies by Fa Wang and collaborators[25, 50, 51] and in subsequent studies[52]. A typical result for the

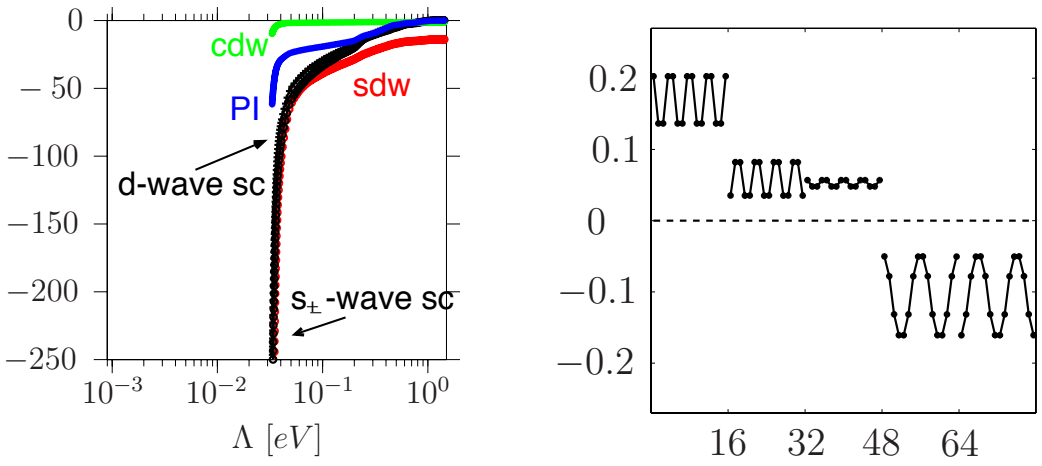


Fig. 5: *N-patch fRG results for the 5-pocket scenario for FeAs-compounds, as laid out in Ref. [52]. Left plot: fRG flow of the leading ordering tendencies. The sign-changing s-wave channel is competing with SDW and d-wave pairing. Right plot: Leading eigenvector of the Cooper pair scattering around the Fermi pockets, showing the sign change between hole- (patch indices 1 to 48) and electron-pockets (indices 49 to 80) and the anisotropy of the suggested gap function. Data taken from Ref. [52], interaction parameters $U = 3.5eV$, $U' = 2.0eV$, $J_H = J_{\rm pair} = 0.7eV$.*

predicted pairing gaps are shown in the right plot of Fig. 5. If the orbital makeup in Eq. 48is set to unity, the gap structure becomes isotropic around the pockets. Another important aspect where the orbital composition is of relevance is the scattering with the wavevector corresponding to the ordering wave vector of the SDW state. Here, initial and final scattering states have differing orbital content, which severely reduces the scattering due to intra-orbital interactions. As an effect, the SDW ordering tendency is weaker than one would estimate based on the pure dispersion nesting. A related version of such orbital mismatch effects is nicely demonstrated in a paper by Kiesel and Thomale[53].
Summarizing this brief section, the iron superconductors pose an interesting problem to the functional RG where the main ordering tendencies have been calculated in good agreement with experiments, at least according to the currently accepted picture. Many more works in this direction are, e.g., mentioned in the recent review by Platt et al. [6]. For the future research, one goal should be to make the fRG a useful bridge between ab-initio descriptions and experimental observables, in particular regarding materials trends in, e.g., the gap structure or the energy scales of the different systems. Furthermore, the studies should be extended to include the dispersion orthogonal to the iron-pnictide planes, as this would yield additional possibilities for nodes in the gap function[1].

5.3 Graphene

Here we describe another string of works where the fRG is used to explore possible ground state ordering in the field of graphene. Single-layer graphene boasts a verity of remarkable phenomena, like the half-integer quantum Hall effect or Klein tunneling[54, 55, 56]. These phenomena can be understood in simplified models without direct inclusion of the electron-electron interaction. Furthermore, at least according to the present state of sample preparation and experimental analysis, the ground state of single-layer graphene is ungapped and there is no spontaneous symmetry breaking even at lowest temperatures. This is usually understood as

an effect of the vanishing electronic density of states at the charge neutrality point (i.e. undoped graphed without external bias). This means that a nonzero critical interaction strength is required to cause any spontaneous symmetry breaking, and the interaction values in graphene are simply below this threshold such that the symmetry breaking does not occur. Doped single-layer graphene may however show interesting many-body instabilities, see e.g. Refs. [57, 58, 59].
Recently, a number of experiments have addressed the ground state properties of bilayer graphene[3] in the Bernal-(AB)-stacking, trying to clarify possible signatures for the interaction between the electrons. Also trilayer systems have been studied[4]. So far, however, the precise nature of the ground state and whether the electronic spectrum is gapped or gapless remain under controversial discussions. In theory, a number of competing instabilities of the semi-metal, e.g. toward a gapped antiferromagnetic (AF) state, a gapless nematic phase and even toward gapped topological phases as the quantum anomalous and quantum spin hall states are debated and compared to the experimental findings[60]. All these theoretical studies indicate that the ground state properties may depend decisively on the profile of the interaction. In this context it is important to notice that the effective interaction parameters and their spatial dependence for the usual low-energy models of graphene and graphite have been calculated by ab-initio techniques[61] using the constrained random phase approximation (cRPA) that takes into account the screening due to bands further away from the Fermi level.
In some recent works we have used a functional RG (fRG) scheme for an unbiased investigation of the instabilities of the Bernal stacked bilayer honeycomb lattice[62, 63], and also on the trilayer[64]. Our approach can be regarded as quantitative, as we integrate out the four p_z-derived bands fully from the band edges toward the Fermi energy and as we use the ab-initio interaction parameters given in Ref. [61]. This way we strongly reduce the variety of possibilities depending on the model parameters and get closer toward a realistic picture of bilayer graphene. We start with a tight-binding model for the Bernal stacked graphene bilayer at the charge neutrality point, as, e.g., given in Refs. [2] or [61]. Within both layers, we include a nearest-neighbor hopping with amplitude t. The interlayer hopping t_p is defined between the sites of the sublattices A in both layers that lie on top of each other. Diagonalizing the tight-binding Hamiltonian results in 4 bands. Two of them have two inequivalent quadratic band crossing points (BCP) K and K' at the Brillouin zone (BZ) corners at the Fermi level. The nonzero density of states ate the Fermi level allows interaction-driven instabilities for arbitrarily small couplings. We ignore trigonal warping terms for the time being, but comment later on the effect of perturbations on the quadratic BCPs. As interactions, we account for an onsite repulsion U, a nearest neighbor density-density interaction V_1 and a second-nearest neighbor density-density interaction V_2. For these terms, cRPA values are listed in Ref. [61]. In addition, for checking the robustness of our results, we considered a third-nearest neighbor repulsion V_3 as well.
For the detection of instabilities, we employ a functional renormalization group (fRG) approach for the one-particle-irreducible vertices of a fermionic many-body system described in the sections before. The discretization of the interaction vertex V is implemented by dividing the BZ into N patches with constant wavevector-dependence within one patch. The representative momenta for the patches are chosen to lie at or close to the Fermi level, see Fig. 6. Each of the four momenta in the interaction vertex is additionally equipped with a sublattice index and a layer index. Momentum conservation fixes one of the four wavevectors resulting in a $(4^4 \times N^3)$-component coupling function V.
Here we present fRG results at zero temperature for interlayer hopping $t_p = 0.1t$. The next-neighbor hopping $t \sim 2.8$eV sets the energy unit. We study the parameter space spanned by

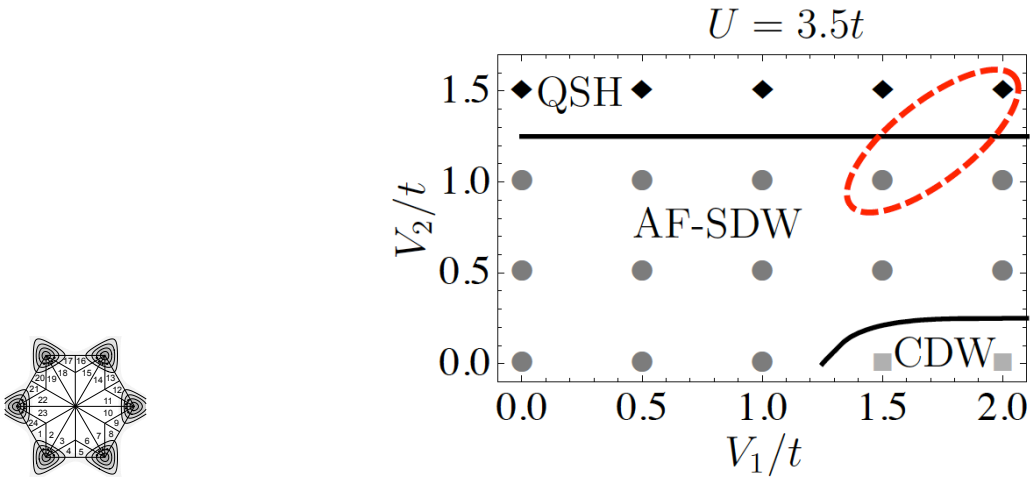


Fig. 6: *Left plot: Patching scheme of the bilayer Brillouin zone in the fRG. The black dots denote the momentum vectors at which the coupling function is evaluated. Right: Tentative phase diagrams for the leading ordering as obtained from fRG as function of the nearest neighbor (V_1) and next-nearest neighbor (V_2) density-density repulsion. The red ellipse indicates the cRPA value for the interaction parameters for single-layer graphene (upper right corner of ellipse) and graphite (lower left corner).*

U, V_1 and V_2 up to the cRPA parameters in Ref. [61] We find flows to strong coupling with non-zero critical scales for all choices of non-vanishing interaction terms. By identifying the leading tendencies, i.e. the strongest class of divergent couplings, we encounter rich tentative phase diagrams as shown in Fig. 6. The observed phases are discussed in the following.

a) An antiferromagnetic (AF) spin density wave (SDW) instability. In the fRG, the flow towards the AF-SDW is seen as a leading divergence of interaction components with zero momentum transfer in the spin channel with AF sublattice structure. The effective interaction is infinitely-ranged due to its sharpness in momentum space. A mean-field decoupling results in an AF spin alignment in each layer where a net spin (e.g. 'up') moment is located on the A1- and B2-sublattices, and an opposite net spin ('down') moment on the B1- and A2-sublattices.

b) A charge density wave (CDW) instability. Here we encounter diverging interactions in the density channel, again with zero momentum transfer. The sign-structure between the layers favors an enhanced occupancy of the A1- and B2-sublattices and a reduced occupancy on the B1 and A2 sublattices or vice versa. The electronic spectrum is gapped.

c) A quantum spin Hall (QSH) instability. This phase breaks spin-rotational symmetry, whereas time reversal symmetry is conserved. In the fRG, spin interactions with zero wavevector transfer diverge, with a sign structure that alternates between K and K' points, and between the sublattices. In a mean-field treatment, a purely imaginary Kane-Mele[65] order parameter is induced. In the fRG, the chirality of the state comes out the same in the two layers for the same spin, i.e. there should be two edge modes with the same propagation direction per spin. Hence, the edge states would not be topologically protected, in contrast with the single-layer case[66].

We found another 'CDW3' instability for smaller U and $V_2/t < 1$, which was until then unmentioned in the literature. Here a site-centered CDW tendency with a finite momentum transfer $Q = K - K' = K'$ grows during the flow. In a variational treatment, the corresponding effective interaction is minimized by complex expectation values. These break the translational symmetry by density modulations. Each sublattice is broken up into three, hence the name CDW3. This

state should be observable directly in scanning tunneling experiments. The fermionic spectrum is gapless and has 4 Dirac cones near the center of the new reduced BZ.
While the first two states, SDW and CDW, are relatively conventional orderings, and the use of the fRG mainly consists in describing the competition between them, the QSH state and the CDW3 are rather exotic. These examples nicely demonstrate that instabilities of systems with more than one band or one state per unit cell offer a greater variety of possibilities. The virtue of the fRG is that it can tell which of the many possibilities is most likely to occur.
Now we discuss the relation to bilayer-graphene. In Ref. [61] the interaction parameters for single-layer graphene and graphite were estimated by ab-initio methods. We expect bilayer graphene to interpolate between these cases. The area of these ab-initio interaction parameters is shown as a red dashed line in Fig. 6. In the fRG for this parameter range SDW and QSH instabilities compete. As both states have a non-zero single-particle-gap, they are compatible with the gapped experimental spectra of Refs. [3]. We also investigated the impact of non-vanishing chemical potentials to mimic the effect of impurities or small dopings on the groundstate, with focus on the cRPA interaction parameters and $t_p = 0.1t$. In the range between for μ from 0 to $0.5t$, the critical scale only changes mildly and the groundstate remains unchanged.
Another issue is that the theoretical instability scales deduced for these realistic parameters are huge (up to $\sim t$), due to the perfect nesting between the two bands forming the quadratic BCPs. Instead, the experimental energy gaps in Refs. [3] are ~ 2 meV $\sim 10^{-3}t$ and hence orders of magnitude smaller than the gaps one gets theoretically in this simple modeling. This poses an important question for further research. We cannot guarantee that our method still works quantitatively in this regime. However, a recent QMC analysis confirms our fRG result of high critical scales[63]. Hence, a quantitative understanding remains to be found. Note that as the instability scales are larger than the usually expected perturbations of the quadratic dispersion near the K, K' points, these additional effects should not resolve this problem. It appears however likely that a combination of theoretical overestimations and imperfections and deviations in the experimental systems must be taken in order to account for the discrepancy.
In summary, we have desrcibed fRG studies of interaction driven instabilities in the honeycomb bilayer model. Besides a novel gapless CDW state, we found that using ab-initio estimates for the band-structure and non-local interaction parameters for bilayer graphene leads to a narrow competition of quantum-spin-Hall and AF-SDW instabilites, making them the two main candidates for the experimental search. The results for the ABC-stacked trilayer systems are quite similar (see Ref. [64]). Details might decide what the actual groundstate is. Important quantitative information from our study is the high energy scale for the breakdown of the gapless state. At present, these high scales are not reflected in the experiments, and more research is needed to understand this discrepancy.

6 Conclusions

The multi-patch fRG approach to interacting electron systems on two-dimensional lattices has provided a number of interesting physical insights. In one-band cases the fRG mainly served as a constancy check for previous other weak coupling approaches, and as an extension of simpler perturbative techniques such as low-order expansions, RPA or spin-fluctuation theories. The fRG serves as an important check for the existence of a d-wave paired ground states in the one-band Hubbard model on the square lattice. In multi-band models, the wave vector resolution of the fRG has allowed one to identify additional aspects and their reasons, such as

the gap anisotropy in iron pnictides and the important role of the orbital character of the band in determining instability scales. In graphene systems, the fRG helps to disentangle the larger number of suggested ground state orderings.
The next step should now be to make these methods more quantitative such they can be used in more material-specific questions. Various groups have already started to combine ab-initio explorations of the band structure and interaction parameters in Wannier bases with the fRG instability search[6, 22, 62, 64]. Here the fRG can be expected to go beyond mean-field, RPA or other standard approaches to explore the consequences of the interactions at lower energy scales. The goal should be to improve the fRG scheme to deliver realistic numbers for gap scales and parameter trends. Then the scheme may be directly useful for the deliberate design of functional materials.

References

[1] An early review is M.R. Norman Physics **1**, 21 (2008), see also K. Ishida, Y. Nakai, and H. Hosono, J. Phys. Soc. Jpn. **78**, 062001 (2009) for a later experimental review; P. J. Hirschfeld, D. J. Scalapino, Physics **3**, 64 (2010) for a recent discussion of the pairing state.

[2] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim, Rev. Mod. Phys. **81**, 109 (2009).

[3] See, e.g., J. Velasco Jr. et al., Nature Nanotechnology 7, 156 (2012); F. Freitag et al., Phys. Rev. Lett. **108**, 076602 (2012); R. T. Weitz, M. T. Allen, B. E. Feldman, J. Martin, and A. Yacoby, Science **330**, 812 (2010).

[4] W. Bao et al., Nature Physics **7**, 948 (2011).

[5] Walter Metzner, Manfred Salmhofer, Carsten Honerkamp, Volker Meden, and Kurt Schönhammer, Rev. Mod. Phys. **84**, 299 (2012).

[6] C. Platt, W. Hanke, R. Thomale, to appear in Adv. Physics 2013.

[7] P. Kopietz, L. Bartosch, F. Schütz, *Introduction to the Functional Renormalization Group*, Springer Lecture Notes in Physics 798, Springer 2010.

[8] C. Honerkamp, Chapter 12 in *Strongly Correlated Systems, Theoretical Methods*, Eds. Adolfo Avella, Ferdinando Mancini, Springer Series in Solid-State Sciences, Volume 171 (2012).

[9] J. Reuther and P. Wölfle, Phys. Rev. B **81**, 144410 (2010); J. Reuther and R. Thomale, Phys. Rev. B **83**, 024402 (2011); J. Reuther, P. Wölfle, R. Darradi, W. Brenig, M. Arlego, and J. Richter, Phys. Rev. B **83**, 064416 (2011).

[10] See textbooks like J.W. Negele, H. Orland, *Quantum Many-particle systems*, Advanced Book Classics, Westview Press.

[11] F. Aryasetiawan, M. Imada, A. Georges, G. Kotliar, S. Biermann, and A. I. Lichtenstein, Phys. Rev. B **70**, 195104 (2004); Takashi Miyake and F. Aryasetiawan, Phys. Rev. B **77**,

085122 (2008); Takashi Miyake, Ferdi Aryasetiawan, and Masatoshi Imada, Phys. Rev. B **80**, 155134 (2009).

[12] M. Salmhofer, C. Honerkamp, Prog. Theor. Phys. **105**, 1 (2001).

[13] K.-U. Giering and M. Salmhofer, Phys. Rev. B **86**, 245122 (2012); C. Husemann, K.-U. Giering, and M. Salmhofer, Phys. Rev. B **85** 075121 (2012).

[14] S. Uebelacker, C. Honerkamp, Phys. Rev. B **86**, 235140 (2012).

[15] S.A. Maier, C. Honerkamp, Phys. Rev. B **85**, 064520 (2012).

[16] A. A. Katanin, Phys. Rev. B **70**, 115109 (2004).

[17] F. Schütz, L. Bartosch, P. Kopietz, Phys. Rev. B **72**, 035107 (2005).

[18] S. G. Jakobs, V. Meden, H. Schoeller, Phys. Rev. Lett. **99**, 150603 (2007).

[19] C. Honerkamp, M. Salmhofer, Phys. Rev. B **64**, 184516 (2001); Phys. Rev. Lett. **87**, 187004 (2001).

[20] C. Honerkamp, D. Rohe, S. Andergassen, T. Enss, Phys. Rev. B **70**, 235115 (2004).

[21] C. Husemann and M. Salmhofer, Phys. Rev. B **79**, 195125 (2009).

[22] W.-S. Wang, Y.-Y. Xiang, Q.-H. Wang, F. Wang, F. Yang, and D.-H. Lee, Phys. Rev. B **85**, 035414 (2012); W.-S. Wang, Z.-Z. Li, Y.-Y. Xiang, and Q.-H. Wang, Phys. Rev. B **87**, 115135 (2013).

[23] D. Zanchi, H. J. Schulz, Europhys. Lett. **44**, 235 (1998); D. Zanchi, H.J. Schulz, Phys. Rev. B. **61**, 13609 (2000).

[24] S. Maier, J. Ortloff, C. Honerkamp, Phys. Rev. B **88**, 235112 (2013).

[25] F. Wang, H. Zhai, Y. Ran, A. Vishwanath, and D.-H. Lee, Phys. Rev. Lett. **102**, 047005 (2009).

[26] C. Karrasch, R. Hedden, R. Peters, Th. Pruschke, K. Schönhammer, V. Meden, J. Phys.: Condensed Matter **20**, 345205 (2008).

[27] F. C. Zhang, T. M. Rice, Phys. Rev. B **37**, 3759 (1988).

[28] C. Proust, E. Boaknin, R. W. Hill, L. Taillefer, A. P. Mackenzie, Phys. Rev. Lett. **89**, 147003 (2002).

[29] M. Abdel-Jawad, M. P. Kennett, L. Balicas, A. Carrington, A. P. Mackenzie, R. H. McKenzie, N. E. Hussey, Nature Physics **2**, 821 (2006); M. Abdel-Jawad, J. G. Analytis, L. Balicas, A. Carrington, J. P. H. Charmant, M. M. J. French, A. P. Mackenzie, and N. E. Hussey, Phys. Rev. Lett. **99**, 107002 (2007).

[30] T. Baier, E. Bick, C. Wetterich, Phys. Rev. B **70**, 125111 (2004).

[31] C. J. Halboth, W. Metzner, Phys. Rev. B **61**, 7364 (2000); Phys. Rev. Lett. **85**, 5162 (2000).

[32] C. Honerkamp, M. Salmhofer, N. Furukawa, T. M. Rice, Phys. Rev. B **63**, 035109 (2001).

[33] S.-W. Tsai, J. B. Marston, Can. J. Phys. **79**,1463 (2001) and arXiv:cond-mat/0010300 (unpublished).

[34] H.H. Lin, L. Balents, M.P.A. Fisher, Phys. Rev. B **56**, 6569 (1997); H.H. Lin, L. Balents, M.P.A. Fisher, Phys. Rev. B **58**, 1794 (1998); M.P.A. Fisher in *Topological Aspects of Low Dimensional Field Theories*, Les Houches Lecture Session LXIX, Eds. A. Comtet, T. Jolicoeur, S. Ouvry, F. David (Springer, 1999).

[35] N. Furukawa, T.M. Rice, M. Salmhofer, Phys. Rev. Lett. **81**, 3195 (1998).

[36] A. Läuchli, C. Honerkamp, T.M. Rice, Phys. Rev. Lett. **92**, 037006 (2004).

[37] M. Salmhofer, C. Honerkamp, W. Metzner, and O. Lauscher, Prog. Theor. Phys. 112, 943 (2004).

[38] Andreas Eberlein, Walter Metzner, arXiv:1308.4845.

[39] C. Honerkamp, Eur. Phys. J. B **21**, 81 (2001).

[40] V. I. Anisimov, Dm. M. Korotin, M. A. Korotin, A. V. Kozhevnikov, J. Kunes, A. O. Shorikov, S. L. Skornyakov, S. V. Streltsov, J. Phys. Condens. Mat. **21**, (2009) 075602.

[41] T. Miyake, K. Nakamura, R. Arita, M. Imada, J. Phys. Soc. Jpn. **79**, 044705 (2010).

[42] I. I. Mazin, D. J. Singh, M. D. Johannes, M. H. Du, Phys. Rev. Lett. **101**, 057003 (2008).

[43] M. Daghofer, A. Nicholson, A. Moreo, E. Dagotto, Phys. Rev. B **81**, 014511 (2010).

[44] K. Kuroki, S. Onari, Ryotaro Arita, Hidetomo Usui, Yukio Tanaka, Hiroshi Kontani, Hideo Aoki, Phys. Rev. Lett. **101**, 087004 (2008).

[45] H. Ikeda, R. Arita, J. Kunes, Phys. Rev. B **81**, 054502 (2010).

[46] A.V. Chubukov, D.V. Efremov, and I. Eremin, Phys. Rev. B **78**, 134512 (2008).

[47] Ch. Platt, C. Honerkamp and W. Hanke, New J. Phys. **11**, 055058 (2009).

[48] T.A. Maier, S. Graser, D.J. Scalapino, P.J. Hirschfeld, Phys. Rev. B **79**, 224510 (2009); S. Graser, T. A. Maier, P. J. Hirschfeld, D. J. Scalapino, New J. Phys. **11**, 025016 (2009).

[49] K. Kuroki, H. Usui, S. Onari, R. Arita, H. Aoki, Phys. Rev **101**, 087004 (2008).

[50] F. Wang, H. Zhai, D.-H. Lee, Europhys. Lett., **85**, 37005 (2009).

[51] F. Wang, H. Zhai, D-H. Lee, Phys. Rev. B **81**, 184512 (2010).

[52] R. Thomale, Ch. Platt, W. Hanke, B. A. Bernevig, Phys. Rev. Lett. **106**, 187003 (2011).

[53] M. Kiesel and R. Thomale, Phys. Rev. B **86**, 121105 (2012).

[54] Novoselov, K. S., Geim, A. K., Morosov, S. V., Jiang, D., Katsnelson, M. I., Grigorieva, I. V., Dubonos, S. V., and Firsov, A. A., Nature **438**, 197 (2005); Zhang, Y., Tan, Y.-W., Stormer, H. L., and Kim, P., Nature **438**, 201 (2005).

[55] Katsnelson, M. I., K. S. Novoselov, and A. K. Geim, Nat. Phys. **2**, 620 (2006).

[56] A.K. Geim and K.S. Novoselov, Nature Materials **6**, 183 (2007).

[57] M. L. Kiesel, C. Platt, W. Hanke, D. A. Abanin, and R. Thomale, Phys. Rev. B **86**, 020507 (2012).

[58] W.-S. Wang, Y.-Y. Xiang, Q.-H. Wang, F. Wang, F. Yang, and D.-H. Lee, Phys. Rev. B **85**, 035414 (2012).

[59] Wei Wu, Michael M. Scherer, Carsten Honerkamp, Karyn Le Hur, Phys. Rev. B **87**, 094521 (2013).

[60] E. V. Castro, N. M. R. Peres, T. Stauber, N. A. P. Silva, Phys. Rev. Lett. **100**, 186803 (2008); R. Nandkishore, L. Levitov, Phys. Rev. Lett. **104**, 156803 (2010); F. Zhang, H. Min, M. Polini, A. H. MacDonald, Phys. Rev. B **81**, 041402 (2010); Y. Lemonik, I. L. Aleiner, C. Toke, V. I. Fal'ko, Phys. Rev. B **82**, 201408 (2010); O. Vafek, K. Yang, Phys. Rev. B **81**, 041401 (2010); Yong Wang, Hao Wang, Jin-Hua Gao, Fu-Chun Zhang, Phys. Rev. B **87**, 195413 (2013); and many other works.

[61] T. O. Wehling et al., Phys. Rev. Lett. **106**, 236805, (2011).

[62] Michael M. Scherer, Stefan Uebelacker, and Carsten Honerkamp, Phys. Rev. B **85**, 235408 (2012).

[63] Thomas C. Lang, Zi Yang Meng, Michael M. Scherer, Stefan Uebelacker, Fakher F. Assaad, Alejandro Muramatsu, Carsten Honerkamp, and Stefan Wessel, Phys. Rev. Lett. **109**, 126402 (2012).

[64] Michael M. Scherer, Stefan Uebelacker, Daniel D. Scherer, and Carsten Honerkamp, Phys. Rev. B **86**, 155415 (2012).

[65] C. L. Kane and E. J. Mele, Phys. Rev. Lett. **95**, 226801 (2005); a related spinless version by F. D. M Haldane, Phys. Rev. Lett. **61**, 2015 (1988).

[66] S. Raghu, Xiao-Liang Qi, C. Honerkamp, and Shou-Cheng Zhang, Phys. Rev. Lett. **100**, 156401 (2008).

C 6 Theory of Scanning Tunneling Microscopy[1]

S. Lounis
Peter Grünberg Institut & Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Scanning tunneling microscopy (STM) is a tool that profoundly shaped nanoscience and nanotechnology. Since its invention by Rohrer and Binnig [1, 2, 3], for which they were awarded the 1986 Nobel prize in Physics, STM experienced revolutionary developments allowing to see for the first time nanostructures at the atomic scale. Another one is to access spintronics at the nanolevel. With increasing availability of low-temperature STM, local electronic properties can be investigated with unprecedented space and energy resolution which opens the vista to completely new applications. STM allowed the rather unique ability of accessing at the same time occupied and unoccupied electronic states. In Fig. 1 is shown a schematic view of the chronological achievements of STM during the last 20 years. Although one cannot mention all important milestones in a single figure, Fig. 1 tells us that after the initial application of STM as a visualization tool of substrates at the atomic level (surface topography), it developed quickly into a device for the manipulation of atoms. Indeed in the 90's, nanostructures such as corrals were built atom by atom whereby a fundamental quantum property, Friedel oscillations induced by the presence of impurities in an electron gas, were observed and confined within man-made nano-objects [4, 5, 6, 7]. These achievements were the prelude to functionalization of nanostructures for different applications, with the aim of characterizing and manipulating not only the spin and charge of single atoms or single molecules but also their position in much bigger nanostructures.

At the beginning of the 21^{st} century, spin-polarized STM (SP-STM) was invented [8, 9] and applied for the investigation of magnetic layers on different substrates. It was, for example, found that a manganese monolayer deposited on tungsten(110) substrate is characterized by an antiferromagnetic ground state, confirming previous predictions made with first-principles calculations [10], and it was nicely shown that contrary to the regular STM, the spin-polarized version shows a magnetic superstructure on top of the chemical unit-cell. Magnetic characterization is nowadays a routine work, that allowed the discovery of new magnetic states, for example chirality, induced by the existence of Dzyaloshinskii-Moriya interactions, was revealed in the magnetic contrast measured on a manganese layer deposited on W(110) surface [11] (See Chapter **C 4** by S. Blügel). Moreover, magnetism of finite nanostructures, nanowires for instance, on magnetic surfaces was recently characterized. [12, 13]

Being the ubiquitous apparatus for nanoscience, the range of phenomena studied by STM is continuously growing. Besides surface topography, and the investigation of ground state properties of excitations, vibrational [14], magnetic [15, 16, 17, 18] or optical [19] properties, which allow chemical identification of atoms, and even measurement of their magnetic anisotropy energy is today a major topic studied with state-of-the-art machines. Recently, magnetometry measurements allowed to extract quantitative values for magnetic exchange interactions among adatoms separated by large distances! [20, 21] Also other applications and developments of STM are geared towards the measurements of adhesion and strength of individual chemical bonds, friction, studies of dielectric properties, contact charging, molecular manipulation and many other phenomena from the micrometer down to subnanometer scale. As Chen says in his book [22]: *It was often said that STM is to nanotechnology what the telescope was to astronomy. Yet STM is capable of manipulating the objects it observes, to build nanoscale structures never existed in Nature. No telescope is capable of bringing Mars and Venus together.*

The actual playground of STM experiments was initially covered mainly by theory. The advent of such an instrument urged the theoretical community for the development of new methods that allows the understanding and prediction of phenomena accessible with STM (see e.g. Refs. [23,

24, 25, 26, 27] or Refs. [28, 29, 30, 31, 32, 33, 34, 35] and many others). It is clear that this tool will continue to experience further evolutions and to play a pivotal role in further developments of nanosciences. Thus more challenges will be proposed to theoreticians. It is not a surprise if after 20 years from its inventions, several books and reviews were dedicated to this technique, to cite a few see Refs. [36, 37, 38, 39, 40, 41, 42, 43].
The goal of this lecture is to review the basics behind the theory accompanying the experiment which could be of interest for readers aiming to work in this field or for those who want to reassess some of the fundamental concepts. Several flavors of the STM method have been developed and invented, we cannot go over all of them but we will discuss the standard ones following partly the book of Chen [22] and the lecture notes of Stefan Blügel [44].

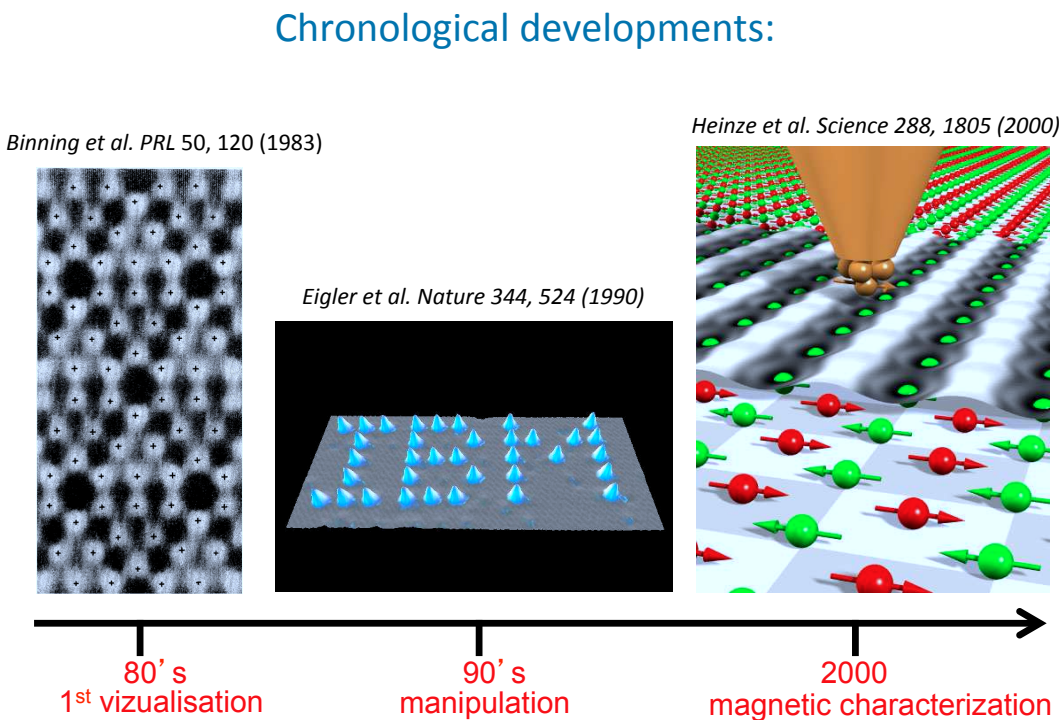


Fig. 1: *Chronological developments of STM. At the early days of its invention, STM was used as a visualization tool for surface topography. It developed into a unique tool that allows access to occupied and unoccupied electronic states. Then atomic manipulation was achieved and magnetic contrasts were realized thanks to the spin-polarized STM. We note that it is impossible to mention all milestones that STM allowed to reach.*

2 Description of scanning tunneling microscopy

Before discussing the basic theory explaining the measurements that can be done with STM, a description of this tool is needed. A crucial ingredient in any STM is the probe tip that is attached to a piezodrive, which consists of three mutually perpendicular piezoelectric transducers (x, y and z piezo). Upon applying a voltage, a piezoelectric transducers contracts or expands which allows to move the tip on the surface. Since the tip is not touching the substrate, the flowing current, I, is weak and is obtained via a tunneling mechanism through the vacuum.

Fig. 2(a) shows a schematic picture of an STM device. If a bias voltage, V, is applied between the tip and the sample, the tunneling current can change.

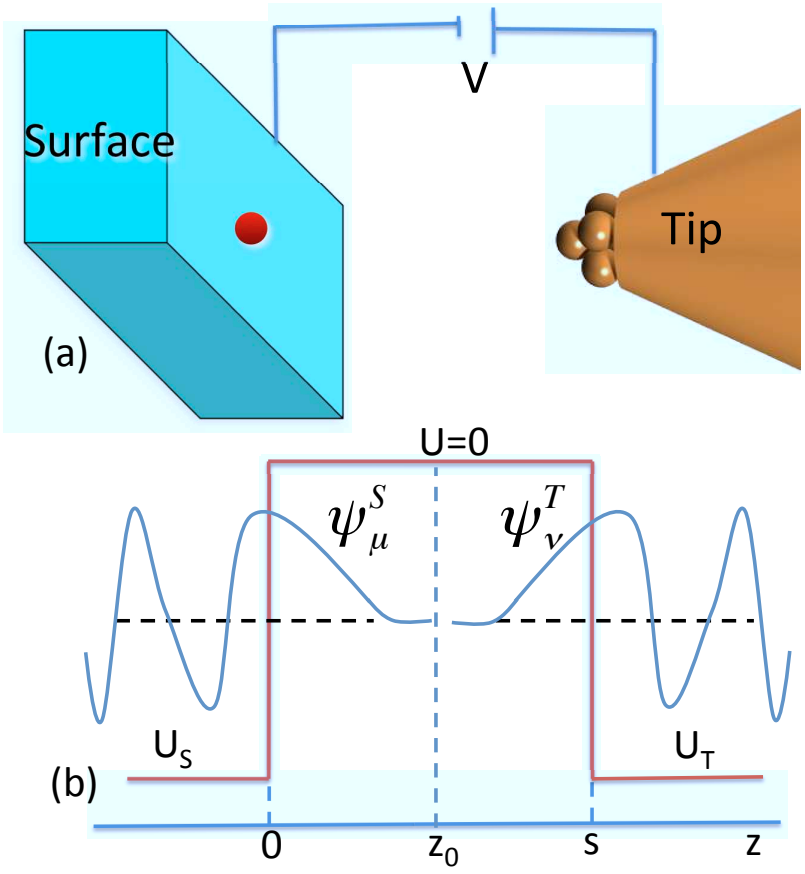


Fig. 2: *A schematic picture of STM is shown in (a). A current can flow between the tip and the substrate through vacuum via a tunneling mechanism. In quantum mechanics, a particle has a non-zero probability of tunneling through a potential barrier which in the STM case is induced by the vacuum. A simple barrier as shown in (b) explains the physics of tunneling. When the two electrodes are far apart, the wave functions of both electrodes A and B decay into the vacuum while the tunneling can take place if the electrodes are closer.*

The simplest way to obtain a scanning tunneling microscope image is to directly measure the variation of the tunnel current as a function of the scanning position while keeping the distance between tip and sample surface constant. A so-called current image is then obtained. Instead of directly recording the atomic variation of the current, however the usual procedure is to keep the tunnel current constant while scanning over the surface. This is done by changing the distance between the tip and surface using a feedback loop. In order to get an image, the voltage required at the piezoelectric crystal to adjust the distance is recorded. One obtains then the so-called constant-current STM image. A further operation mode is the spectroscopy acquisition by STM. It is usually done by interrupting the feedback in order to keep for the I–V spectroscopy data acquisition the tip-sample separation constant. This can be done at any

desired surface spot or for every pixel in a STM image. An extended discussion of the different operating modes will be given once the theory of STM is presented.
Obviously, if one wants to understand the working mechanism of STM and simulate the experiment, one could think of simulating the whole setup, i.e. considering a tip, a substrate, a bias voltage and calculate the tunneling current or conductance (see e.g. Chapter **A 9** by S. Tsukamoto). Although actual ab-initio methods are capable of handling few hundreds up to few thousands of atoms in a unit-cell (see e.g. Chapter **D 6** by R. Zeller), technical issues can occur. For instance, a periodic supercell approach would lead to a non-realistic scenario of multiples tips scanning the substrate at the same time. Methods based on Green functions allow to consider two perfect semi-infinite substrates separated by vacuum. One of the substrates would simulate the surface, on the other substrate, a model tip can be embedded (see e.g. Ref. [45]). Although this scheme is appealing, one would be facing the problem of choosing the right model for the tip, which is far from being easily accessible experimentally. All of those arguments stimulated approximations that are very often used successfully for the understanding of STM-experiments but they also bear limitations.

3 The concept of tunneling

Here we describe briefly elementary theories of tunneling through a one-dimensional potential barrier, which will help us to grasp the basic concept used in STM. In quantum mechanics, the electron feeling a potential U(z), for example the one shown in Fig. 2(b) considering $U_S = U_T = -U$, is described by a wave function $\psi(z)$, which satisfies the Schrödinger equation,

$$-\frac{\hbar^2}{2m}\frac{d^2}{dz^2}\psi(z) + U(z)\psi(z) = E\psi(z), \tag{1}$$

at a given position z. In this elementary model, the STM setup is simplified to a one-dimensional potential barrier where the vacuum is modeled by the potential barrier U, while its left and right sides shown in Fig. 2(b) represent the substrate, S, and the tip, T. When $E > |U|$, the solutions of Eq. 1 are $\psi(z) = \psi(0)e^{\pm ikz}$ where $k = \frac{\sqrt{2m(E-|U|)}}{\hbar}$ is the wave vector. In the classically forbidden region, within the barrier, the solution is given by $\psi(z) = \psi(0)e^{-\varkappa z}$ with $\varkappa = \frac{\sqrt{2m(|U|-E)}}{\hbar}$ being the decay constant that describes an electron penetrating through the barrier into the $+z$ direction. The probability density for the observation of an electron near a point z is finite in the barrier region and is proportional to $|\psi(0)|^2 e^{-2\varkappa z}$. Additionally, electrons can propagate in the opposite direction ($-z$-direction) indicating that tunneling is bidirectional. The total wave function in every region, sample, barrier and tip are written as:

$$\psi_S = e^{ikz} + Ae^{-ikz} \tag{2}$$

$$\psi_{\text{Barrier}} = Be^{-\varkappa z} + Ce^{\varkappa z} \tag{3}$$

$$\psi_T = De^{ikz} \tag{4}$$

The coefficients A, B, C and D take care of the reflection and transmission of the electrons, they are obtained by matching of the wave functions and their derivatives $d\psi/dz$ at the two interfaces, sample–barrier and barrier–tip. The incident current density $I_i = \hbar k/m$ and the transmitted current I_t

$$I_t = -i\frac{\hbar}{2m}\left(\psi_T^*(z)\frac{d\psi_T(z)}{dz} - \psi_T(z)\frac{d\psi_T^*(z)}{dz}\right) = \frac{\hbar k}{m}|D|^2 \tag{5}$$

Element	Al	Au	Cu	Ir	Ni	Pt	Si	W
$\Phi(eV)$	4.1	5.4	4.6	5.6	5.2	5.7	4.8	4.8
$\varkappa(\text{Å}^{-1})$	1.03	1.19	1.09	1.21	1.16	1.22	1.12	1.12

Table 1: *Work functions and decay constants according to Ref. [22] for selected materials.*

define the barrier transmission coefficient T which is given by the ratio between the transmitted current density and the incident current density:

$$T = \frac{I_t}{I_i} = |D|^2 = \frac{1}{1 + \frac{(k^2+\varkappa^2)^2}{4k^2\varkappa^2} \cdot \sinh^2(\varkappa s)} \tag{6}$$

which simplifies in the limit of a strongly attenuating barrier (large decay constant $\varkappa$)

$$T \sim \frac{16\varkappa^2 k^2}{(k^2+\varkappa^2)^2} \cdot e^{-2\varkappa s} \tag{7}$$

where s defines the location of electrode T (Fig. 2(b)).
From this basic model, some important features of a more realistic metal-vacuum-metal tunneling can be explained. Let us first evaluate the decay constant magnitude which is defined defined by the work function Φ primarily if the electrons involved in the tunneling process are lying close to the Fermi energy of both electrodes. Indeed Φ is defined by the minimum energy required to remove an electron from the bulk to the vacuum level. In general, the work function depends not only on the material, but also on the crystallographic orientation of the surface but to simplify our discussion we assume it to be the same for the tip and sample ($\Phi_S = \Phi_T = \Phi$). In our model $|U_S|$ and $|U_T|$ are respectively replaced by their respective work functions. Here the decay constant $\varkappa = \frac{\sqrt{2m\Phi}}{\hbar}$ is of the order of $\sim 1\text{Å}^{-1}$ for the typical values of the work function ($\sim 5eV$). The typical values of work functions of materials used in STM experiments, together with the decay constants, are listed in Table 1. According to Eq. 7, the current decays by one order of magnitude per 1Å.
Even though this model is too simple to describe realistic STM experiments it explains the high sensitivity to height changes in the sample topography. Also it demonstrates that during tunneling, the tip's atom, that is the closest to the substrate, is the main atom involved in the tunneling process!
From this simple one dimensional model one can derive the principle used by Binning and Rohrer when they invented the STM. Their argument to explain the ability of STM to achieve large lateral resolutions and by that probe the electronic structures of various materials at an atomic scale ($\sim$ 2Å) is: because of the tunneling through vacuum, a large lateral resolution much smaller than the radius of the tip-end, R, is possible if the distance between the tip-end and the sample surface, Δz, is much smaller than the tip radius [46]. Near the tip-end, the current lines are almost perpendicular to the sample surface (Fig. 3). At a point Δx on the tip, the distance to the sample surface, z, is increased by

$$\Delta z \sim \frac{\Delta x^2}{2R}. \tag{8}$$

Assuming that at each point the tunneling current density follows the formula for the one-dimensional case, Eq. 7, the lateral current distribution is

$$I(\Delta x) \sim e^{-2\varkappa \frac{\Delta x^2}{2R}}. \tag{9}$$

Typically, $\varkappa \sim 1\text{Å}^{-1}$. For $R = 10\text{Å}$, at $\Delta x \sim 4.5\text{Å}$, the current drops by a factor of $\sim e^{-2}$, that is about one order of magnitude. The diameter of such a current column is the resolution limit, which is about 9Å. Therefore with moderate means, a very high lateral resolution is expected. Nowadays, achievements of STM largely exceeds this expectation.

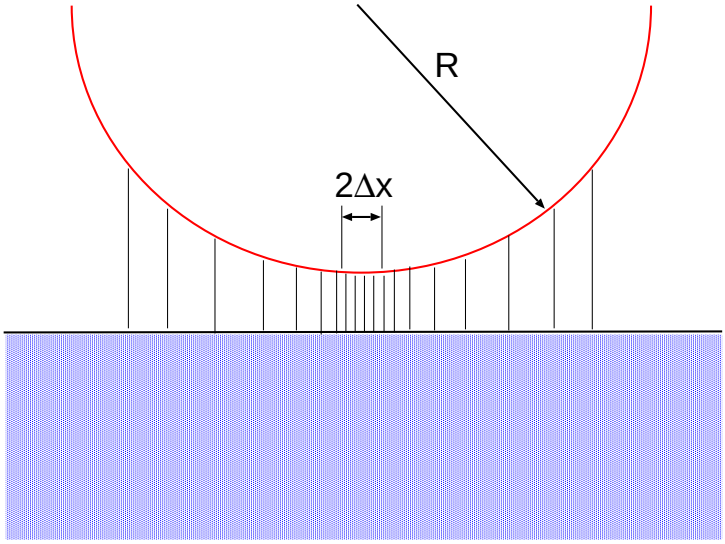


Fig. 3: *Estimation of the lateral resolution in STM. Out of a spherical tip model with radius R very close to the surface, the lateral resolution of STM can be evaluated. The tunneling current is concentrated at the vicinity of the closest point to the substrate.*

4 Modeling currents

4.1 Bardeen's approach

One-dimensional case

The planar tunneling junction problem treated by Bardeen is schematically shown in Fig. 2(b). The model used by Bardeen, called also the transfer Hamiltonian method [47], and extended later on by Tersoff and Hamann [23, 24] and Chen [25, 26] to STM, has naturally limitations by its assumptions but gives a fundamental understanding on the ability of STM to reach high space and energy resolution. Here are some assumptions that are assumed in the original derivation of the Bardeen's approach:
First of all, the electron tunneling is treated as a one-particle process, i.e. the mutual interaction between electrons during tunneling is neglected which is a reasonable approximation in the low tunneling regime. Furthermore, a direct interaction of tip and sample resulting in the formation of coupled electronic states is not taken into account. This assumption is valid if the tip-sample distance is large enough, i.e. a distance larger than ~ 4Å should be sufficient. Note that in our discussion elastic tunneling is assumed, i.e. no energy loss of the electrons with quasi-particles of the electrodes, e.g. plasmons, phonons, spin-excitations is considered in this section. Of course, recently models taking care of this kind of interactions were developed (see Chapter **C 7** by M. Wegewjis for detailed examples) and a simple discussion on the treatment of inelastic tunneling is given later on in this manuscript.

When the two electrodes, representing the tip and the sample, are far apart, the wave functions of electrode S, representing the unperturbed substrate, or of the unperturbed electrode T, representing the tip, satisfy the Schrödinger equation of the free electrode S or T,

$$i\hbar\frac{\partial\psi^i}{\partial t} = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + U_i\right)\psi^i, \tag{10}$$

where U_i is the potential function of electrode i (S or T), and ψ^i depends on both time and spatial coordinates. The stationary states are $\psi^i = \psi^i_\mu e^{-iE^i_\mu t/\hbar}$ with the spatial wave functions and energy eigenvalues satisfying

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + U_i\right)\psi^i_\mu = E^i_\mu\psi^i_\mu, \tag{11}$$

Once the distance between the two electrodes is reduced, the time-evolution of a state ψ in the system tip-sample is governed by the Schrördinger equation containing the full potential:

$$i\hbar\frac{\partial\psi}{\partial t} = \left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + U_S + U_T\right)\psi. \tag{12}$$

The time-evolution can be treated in perturbation theory. At $t \to -\infty$, the tip is far from the substrate and an electron is stationary in a state ψ^S_μ of the sample. We assume that the tip is approached slowly towards the sample and thereby the tip potential is turned on adiabatically. The adiabatic consideration is reasonable since the time-scale of electrons are femtoseconds ($\sim 10^{-15}$sec) while the time needed to move the tip is in seconds. Formally we describe this adiabatic switching of the perturbation via a time-dependent potential

$$U_T(t) = e^{\eta t/\hbar}U_T, \text{ and } \eta > 0. \tag{13}$$

U_T is a constant and η is small and positive. At the end of our derivation, when we consider $\eta \to 0$, the potential will be constant for all times. With the presence of the combined potential, a state ψ^S_μ described by Eq. 11 at $t = -\infty$ will not evolve according to Eq. 10. Instead it has a probability of populating the states of electrode T, denoted as ψ^T_ν. Our goal is to measure that probability since it is directly related to the tunneling current.
The state ψ of the whole system can be expanded in a linear combination of the sample and tip eigenfunctions (as calculated before the perturbation is switched on), which form an orthogonal and complete basis set:

$$\psi = a_\mu(t)\psi^S_\mu e^{-iE^S_\mu t/\hbar} + \sum_{\nu=1}^{\infty} c_\nu(t)\psi^T_\nu e^{-iE^T_\nu t/\hbar}. \tag{14}$$

In our ansatz, $a_\mu(t)$ and $c_\nu(t)$ are coefficients to be determined by Eq. 12 with $a_\mu(-\infty) = 1$ and $c_\nu(-\infty) = 0$. We note that in our ansatz the time evolution coefficients $a_\mu(t)$ and $c_\nu(t)$ is due solely to the presence of the time dependence in $U_T(t)$. As we shall see, this separation is convenient because the time evolution coefficients satisfy a relatively simple differential equation.
It is important to note that each set of wave functions ψ^S_μ and ψ^T_ν originates from different Hamiltonians. Neither of them is an eigenfunction of the Hamiltonian of the combined system. A basic assumption of Bardeen's tunneling theory is that the two sets of wave functions are

approximately orthogonal, $\int \psi_\mu^{T*}\psi_\nu^S d^3r \approx 0$. Inserting Eq. 14 into Eq. 12 and after projection on the state ψ_ν^T, we obtain

$$i\hbar\frac{dc_\nu(t)}{dt} = \left\langle \psi_\nu^T \right| U_T \left| \psi_\mu^S \right\rangle e^{-i\left(E_\mu^S - E_\nu^T + i\eta\right)t/\hbar} + \sum_{\lambda=1}^{\infty} c_\lambda(t) \left\langle \psi_\nu^T \right| U_S \left| \Psi_\lambda^T \right\rangle e^{-i\left(E_\lambda^T - E_\nu^T\right)t/\hbar}. \quad (15)$$

Here we considered the following small approximations: (i) because of the adiabatic approximation we considered the prefactor $a(t)$ to be slowly varying, i.e. $\frac{d}{dt}a_\mu(t) = 0$ and $a_\mu(t) = 1$, and (ii) a second contribution $\sim \left(e^{\eta t/\hbar} - 1\right)$ is neglected since it vanishes at $\eta \to 0$.
This equation can be solved iteratively but we limit ourselves to the first order of time-dependent perturbation theory and neglect the second term on the right-hand side of the previous equation since it is a second-order infinitesimal quantity. Therefore, to first-order,

$$i\hbar\frac{dc_\nu(t)}{dt} = \left\langle \psi_\nu^T \right| U_T \left| \psi_\mu^S \right\rangle e^{-i(E_\mu^S - E_\nu^T + i\eta)t/\hbar}. \quad (16)$$

Since U_T is non-zero only in the volume of electrode T (at $z > s$, see Fig. 2), the integral $\left\langle \psi_\nu^T \right| U_T \left| \psi_\mu^S \right\rangle$, that defines the tunneling matrix element $M_{\mu\nu}$, is evaluated only in the right-hand side of the separation surface. This tunneling matrix element describes the projection of the initial state ψ_μ^S perturbed by the potential U_T onto the final state ψ_ν^T . After integration over time we get

$$c_\nu(t) = \frac{1}{E_\mu - E_\nu + i\eta} M_{\mu\nu} e^{-i(E_\mu^S - E_\nu^T + i\eta)t/\hbar}. \quad (17)$$

$|c_\nu(t)|^2$ describes the probability that an electron initially described by the state ψ_μ^S in time $t = -\infty$ populates a state ψ_ν^T at time t,

$$|c_\nu(t)|^2 = \frac{e^{2\eta t/\hbar}}{\left(E_\mu^S - E_\nu^T\right)^2 + \eta^2} \left|M_{\mu\nu}\right|^2, \quad (18)$$

which leads to the tunneling probability per unit time, $P_{\mu\nu}(t) = \frac{d}{dt}|c_\nu(t)|^2$,

$$P_{\mu\nu}(t) = \frac{2\eta}{\left(E_\mu^S - E_\nu^T\right)^2 + \eta^2} e^{2\eta t/\hbar} \frac{1}{\hbar} \left|M_{\mu\nu}\right|^2, \quad (19)$$

where we can recognize the definition of the $\delta(x)$ distribution given by $\delta(x) = \frac{1}{\pi}\lim_{\eta\to 0}\frac{\eta}{x^2+\eta^2}$. Taking the limit $\eta \to 0$ we find

$$P_{\mu\nu}(t) = \frac{2\pi}{\hbar}\delta\left(E_\mu^S - E_\nu^T\right)|M_{\mu\nu}|^2, \quad (20)$$

which is the famous Fermi's Golden Rule, that is a general result of first order time-dependent perturbation theory (The Golden Rule is also derived in Chapter **A 5** by Ph. Mavropoulos). Elastic tunneling is guaranteed by the delta function $\delta\left(E_\mu^S - E_\nu^T\right)$. The tunneling current I is proportional to $eP_{\mu\nu}$ where e is the elementary electron charge.
Up to now we have considered the tunneling process involving a single state μ to a single state ν. However the tip and substrate are characterized by a continuous spectrum of states, thus we have to consider the sum over states μ and ν for every spin channel. Naturally an electron can only tunnel from an occupied state ψ_μ^S to an unoccupied state ψ_ν^T and vice-versa. At zero temperature, there is a sharp Fermi edge separating occupied and unoccupied states while at

elevated temperatures the Fermi edge is smeared out; occupied states are then described by the Fermi-Dirac distribution $f(E-E_F) = (1+exp\left[(E-E_F)/k_BT\right])^{-1}$ while unoccupied states are described by $1-f(E-E_F)$. Accounting for the occupation in this manner and assuming a bias voltage V, the tunneling current in thermal equilibrium from sample to tip, $I_{S\to T}$, and from tip to sample, $I_{T\to S}$ can be written as:

$$\begin{aligned} I_{S\to T} &= \frac{4\pi e}{\hbar}\sum_{\nu\mu} f(E_\mu^S - E_F^S)\left[1-f(E_\nu^T-E_F^T)\right]|M_{\mu\nu}|^2\,\delta(E_\nu^T-E_\mu^S-eV) \qquad (21)\\ I_{T\to S} &= \frac{4\pi e}{\hbar}\sum_{\nu\mu} f(E_\mu^T - E_F^T)\left[1-f(E_\nu^S-E_F^S)\right]|M_{\mu\nu}|^2\,\delta(E_\nu^T-E_\mu^S-eV) \end{aligned}$$

A factor 2 has been introduced accounting for the two possible spin states of each electron. The difference between the two currents gives a net total tunneling current:

$$I = \frac{4\pi e}{\hbar}\sum_{\nu\mu}\left[f(E_\mu^S-E_F^S) - f(E_\nu^T-E_F^T)\right]|M_{\mu\nu}|^2\,\delta(E_\nu^T-E_\mu^S-eV). \qquad (22)$$

The finite summation over the discrete states can be replaced by an integral over energies using the density of state $n(E)$: $\sum_\mu \to \int n(E)dE$ and after an appropriate change of variable we find

$$\begin{aligned} I &= \frac{4\pi e}{\hbar}\int d\epsilon\left[f(E_F^T-eV+\epsilon) - f(E_F^S+\epsilon)\right] \qquad (23)\\ &\quad \times n^T(E_F^T-eV+\epsilon)n^S(E_F^S+\epsilon)\left|M(E_F^S+\epsilon, E_F^T-eV+\epsilon)\right|^2 \end{aligned}$$

where n^S and n^T are the density of states (DOS) of the substrate and of the tip. We find formally that the tunneling current depends explicitly on the electronic structure of both the tip and substrate which has important consequences on STM measurements. Interestingly at zero temperature or if k_BT is smaller than the energy resolution required in the measurement, the Fermi distribution function can be approximated by a step function and the current simplifies to

$$I = \frac{4\pi e}{\hbar}\int_0^{eV} d\epsilon n^T(E_F^T-eV+\epsilon)n^S(E_F^S+\epsilon)|M|^2 \qquad (24)$$

and for a very small bias voltage

$$I = \frac{4\pi e}{\hbar}Vn^T(E_F^T)n^S(E_F^S)|M|^2. \qquad (25)$$

The differential conductivity, which is the other quantity measured experimentally, is given by

$$\frac{dI}{dV} = \frac{4\pi e}{\hbar}n^T(E_F^T)n^S(E_F^S+eV)|M(E_F^S+eV, E_F^T)|^2. \qquad (26)$$

This explains the unique power of STM to be able to access the occupied and unoccupied electronic states of the substrate. Indeed this can be achieved by changing the sign of the bias voltage V.

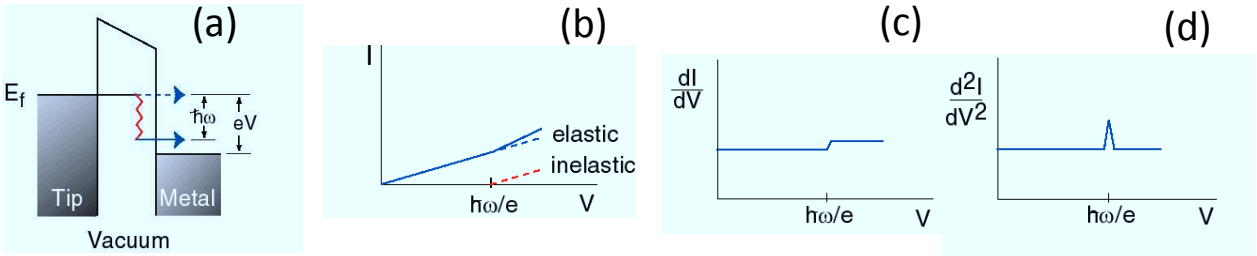


Fig. 4: *Schematic representation of inelastic tunneling with STM (according to Wilson Ho) is shown in (a). If the electron has enough energy, provided by the bias potential, to trigger the excitation mode an additional tunneling channel is created. The slope giving the tunneling current versus the bias voltage (b) changes at a bias voltage corresponding to the frequency of the excitation mode. Taking the first and second derivatives lead to a step-like function (c) or to a resonance (d) at $\hbar\omega$.*

Inelastic tunneling

Although the rest of this lecture is devoted to elastic tunneling phenomena, inelastic tunneling within the Bardeen approach is described briefly in this section. Within STM, these excitations were observed for vibrations (see e.g. [14]), photons (see e.g. [19]) and for spin-excitations (see e.g. [15, 16, 17, 18, 57]).

We have learned earlier that the current and differential conductivity is proportional to the DOS of the substrate and of the tip. Now imagine that on the substrate a molecule is deposited which is characterized by a vibrational mode or a spin excitation mode. If the tunneling current can trigger the excitation, i.e. the tunneling electrons couple to the excitation mode and by that loose their energy, additional tunneling channels can be created. More tunneling possibilities translates to an increase in the tunneling current. Fig. 4 shows schematically how the slope of the current versus the bias voltage increases suddenly at the voltage, or energy, corresponding to the excitation mode. If one calculates or measures the differential conductivity, one obtains step-like functions in the spectra, and a second derivative of the current leads to resonances located at the excitation energies.

Consider that the potential of the sample U_S contains a vibrating adatom and is time-dependent $U_S + U_0\cos(\omega t)$, where U_0 is the amplitude of the vibration and ω is the vibrational frequency of the adatom. We apply once more first order time-dependent perturbation theory as discussed previously and find:

$$i\hbar\frac{dc_\nu(t)}{dt} = M_{\mu\nu}e^{-i(E_\mu^S - E_\nu^T + i\eta)t/\hbar} + \delta M_{\mu\nu}\cos(\omega t)e^{-i(E_\mu^S - E_\nu^T)t/\hbar} \tag{27}$$

where $M_{\mu\nu}$ is the elastic tunneling matrix element $\left\langle \psi_\nu^T \right| U_T \left| \psi_\mu^S \right\rangle$ and $\delta M_{\mu\nu}$ is the inelastic counterpart $\left\langle \psi_\nu^T \right| U_0 \left| \psi_\mu^S \right\rangle$.

After integration, the inelastic contribution to the coefficient c_ν in Eq. 27 is

$$\delta c_\nu(t) = \frac{\delta M_{\mu\nu}}{2}\left[\frac{e^{(E_\mu - E_\nu + \hbar\omega)t/\hbar}}{E_\mu - E_\nu + \hbar\omega} + \frac{e^{(E_\mu - E_\nu - \hbar\omega)t/\hbar}}{E_\mu - E_\nu - \hbar\omega}\right]. \tag{28}$$

The probability is thus simply given by[2]

$$|\delta c_\nu|^2 = |\delta M_{\mu\nu}|^2 \left(\frac{\sin^2[(E_\mu - E_\nu + \hbar\omega)t/2\hbar]}{(E_\mu - E_\nu + \hbar\omega)^2} + \frac{\sin^2[(E_\mu - E_\nu - \hbar\omega)t/2\hbar]}{(E_\mu - E_\nu - \hbar\omega)^2} + \text{ cross terms} \right).$$

The function $\sin^2(xt/2\hbar)/x^2$ reaches its maximum when $x = 0$ and approach rapidly zero for $x \neq 0$. In the limit of long time t, this function is nothing else than a delta function: $t\pi\delta(x)/2\hbar$, which ensures that the inelastic tunneling occurs at $E_\mu - E_\nu - \hbar = \pm\hbar\omega$. Finally we give the probability rate as done in the previous section:

$$\delta P_{\mu\nu} = \frac{d|\delta c_\mu|^2}{dt} \propto \frac{\pi}{2\hbar}|\delta M_{\mu\nu}|^2 \delta(E_\mu - E_\nu \pm \hbar\omega) \tag{29}$$

Therefore in addition to the elastic tunneling term, we have an additional term when the excitation is created, i.e. the energy level of the electronic state changes by an amount $\hbar\omega$. We note that several theoretical methods were developed in order to understand how STM probes excitations. For instance, for spin-excitations, some are based on a Heisenberg Hamiltonian (see e.g. Refs.[29, 30, 31, 32]) and beyond[33], others are based on time-dependent density functional theory [34, 35].

Bardeen's Tunneling Matrix elements

We derived earlier the tunneling current and the differential conductivity in the Bardeen's approach, we investigate in the following the tunneling matrix element $M_{\mu\nu}$. Using Eq. 11, the integral defining the tunneling matrix element $M_{\mu\nu} = \left\langle \psi_\nu^T \right| U_T \left| \psi_\mu^S \right\rangle$ can be converted into a surface integral only depending on the unperturbed wave functions of the two electrodes at the separation surface. By applying Eq. 11, we have

$$M_{\mu\nu} = \int_{z>z_0} \psi_\mu^S \left(E_\nu^T + \frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} \right) \psi_\nu^{T*} d^3r. \tag{30}$$

Because of the elastic tunneling condition, $E_\mu^S = E_\nu^T$, the form giving the tunneling matrix element can be converted into

$$M_{\mu\nu} = \int_{z>z_0} \left(\psi_\mu^S E_\mu^S \psi_\nu^{T*} + \psi_\mu^S \frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2}\psi_\nu^{T*} \right) d^3r. \tag{31}$$

Using $M_{\mu\nu} = \left\langle \psi_\mu^S \right| U_T \left| \psi_\nu^T \right\rangle$ and noticing that, on the tip side, the sample potential U_S is zero, we obtain

$$M_{\mu\nu} = -\frac{\hbar^2}{2m} \int_{z>z_0} \left(\psi_\nu^{T*} \frac{\partial^2 \psi_\mu^S}{\partial z^2} - \psi_\mu^{*S} \frac{\partial^2 \psi_\nu^T}{\partial z^2} \right) d^3r. \tag{32}$$

With the identity

$$\psi_\nu^{T*}\frac{\partial^2\psi_\mu^S}{\partial z^2} - \psi_\mu^S\frac{\partial^2\psi_\nu^{T*}}{\partial z^2} = \frac{\partial}{\partial z}\left[\psi_\nu^{T*}\frac{\partial\psi_\mu^S}{\partial z} - \psi_\mu^S\frac{\partial\psi_\nu^{T*}}{\partial z} \right], \tag{33}$$

the integration over z can be carried out to obtain

$$M_{\mu\nu} = \frac{\hbar^2}{2m}\int_{z=z_0}\left[\psi_\mu^S\frac{\partial\psi_\nu^{T*}}{\partial z} - \psi_\nu^{T*}\frac{\partial\psi_\mu^S}{\partial z} \right] dxdy. \tag{34}$$

[2]Other cross terms are not considered here

The last equation gives Bardeen's tunneling matrix element in a one-dimensional form. It is a surface integral of the wave functions (and its normal derivatives) of the two free electrodes, evaluated at the separation surface. The potential barrier information does not appear explicitly, and only the information of the wave functions at the separation surface is required. Furthermore, the formula is symmetric with regards to both electrodes. It is the basis of the reciprocity principle in STM, which has important consequences in designing and interpreting experimental results.
Although derived for the one-dimensional case, the Bardeen approach can be extended to the three-dimensional case where the tunneling matrix element, Eqs. 32 and 34, change to

$$M_{\mu\nu} = \frac{\hbar^2}{2m} \int_{\Omega_T} \left[\psi_\mu^S \Delta\psi_\nu^{T*} - \psi_\nu^{T*} \Delta\psi_\mu^S\right] \cdot d\vec{r} \tag{35}$$

and

$$M_{\mu\nu} = \frac{\hbar^2}{2m} \int_{\Sigma} \left[\psi_\mu^S \vec{\nabla}\psi_\nu^{T*} - \psi_\nu^{T*} \vec{\nabla}\psi_\mu^S\right] \cdot d\vec{S} \tag{36}$$

where the surface integral is performed on the separation surface, Σ, between the volume defining the sample and the volume defining the tip.

Energy dependence of tunneling matrix elements

The assumption, that the tunneling matrix element M is a constant, is reasonable for a small bias voltage window. However, in scanning tunneling spectroscopy (STS) experiments, the energy scale can be as large as $\pm 2eV$. Thus the energy dependence of the tunneling matrix element cannot be overlooked. The variation of $|M|$ with energy can be evaluated from the Bardeen formula, Eq. 34.
In the gap region, the wave function of electrode S is:

$$\psi_\mu^S(z) = \psi_\mu^S(0) e^{-\varkappa_\mu^S z}, \tag{37}$$

where $\varkappa_\mu^S = \sqrt{2m|E_\mu^S|}/\hbar$ is the decay constant corresponding to the energy eigenvalue of ψ_μ^S. Similarly, in the gap region, the wave function of electrode T is

$$\psi_\nu^T(z) = \psi_\nu^T(s) e^{\varkappa_\nu^T (z-s)}. \tag{38}$$

Because of the condition of elastic tunneling ($E_\mu^S = E_\nu^T$), the two decay constants are equal,

$$\varkappa_\nu^T = \varkappa_\mu^S = \frac{\sqrt{2mE_\mu^S}}{\hbar} \tag{39}$$

Inserting the previous equations into Eq. 34, we obtain

$$M_{\mu\nu} = \frac{\hbar^2}{2m} e^{-\varkappa_\mu^S s} \int_{z=z_0} 2\varkappa_\mu^S \psi_\mu^S(0) \psi_\nu^T(s) dxdy. \tag{40}$$

As expected, the tunneling matrix element is independent of the position of the separation surface, $z = z_0$. The expression in the integral is a constant, because $\psi_\nu^T(s)$ is the value of the wave function of electrode T at its surface. The energy dependence of M is through the decay

constant $\varkappa_\mu^S$. Qualitatively, the effect of the energy dependence of the tunneling matrix element on the tunneling current is as follows. Once the integration over energies is carried out in Eq. 24, we realize that the value of $e^{-\varkappa_\mu^S s}$ near the top of the integral is bigger than its value near the bottom. Therefore, the energy spectrum of electrode T near the Fermi level and the empty states energy from the spectrum of electrode S electrons about eV above the Fermi level are the dominant contributors to the integral in Eq. 24.
We have learned from the Bardeen's approach to calculate the tunneling current, that the exact electronic structure of tip and sample is required. In principle it is possible to calculate them for both systems with the actually available ab-initio methods and to compute all tunneling matrix elements to gain the tunneling current. Although quite elaborate, this scheme is possible. However, the tip structure is not straightforward to access experimentally which complicates the task of simulating the STM tip. This issue pushed the development of models, as the one described in the next section, that simplify the tip's electronic structure. The simplest approach is to get rid off the tip.

4.2 Tersoff-Hamann model

After the invention of STM, Tersoff and Hamann formulated a model [23, 24] based on Bardeen's tunneling theory which is widely used today. Here we describe its concept, derivation and limitations.

The essence of the model

The driving argument behind the Tersoff-Hamann (TH) model is the difficult access to the tip states. Those, as we have learned in the previous sections, are important in the imaging mechanism of STM since the tunneling current is a convolution of electronic states of the tip to those of the sample. Therefore, a particular model of the tip was proposed, such that the tip properties can be simplified and factorized out of the problem. The TH model represents the tip with potential and wave functions arbitrarily localized, in words, modeled as a geometrical point. Consequently, the STM image is related to the properties of the surface alone. Thus, according to that model, STM measures an intrinsic property of the unperturbed surface, rather than a property of the joint surface-tip system.
The TH model has proven to be extremely valuable in interpreting the STM images with characteristic feature sizes of ≥ 10Å, for example, the profiles of superstructures of surface reconstruction, the scattered waves of surface states, as well as defects, adsorbates, and substitution atoms on the surface. However, the TH model predicts that the corrugation of atomic-scale features (with typical length scale close to or smaller than 3Å) is about one picometer or even smaller, which is beyond the detection limit of STM. Also it cannot always explain the rich experimental observations due, obviously, to the convolution of tip electronic states and sample electronic states.

Derivation of TH model

The STM tip is modeled as a locally spherical potential well centered at $\vec{R}_T$. Once more, the sample surface is represented by the $z = 0$ plane. In Bardeen's model the potentials of the tip and sample are negligible in the separation plane Σ. Therefore in the vacuum region, both wave

functions of sample and tip near the Fermi level satisfy the Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta\psi = -\Phi\psi \text{ or } \Delta\psi = \varkappa^2\psi, \tag{41}$$

With the approximation that the tip is just a single atom which has an s-orbital as wave function, this equation has two solutions, an irregular and a regular solution, which are the spherical, modified Bessel functions of first and second kind. We are interested in the regular solution that is characterized by an exponential decay from tip to vacuum. Thus the solution of the previous equation is given by the modified Bessel function of the second kind

$$\psi_\nu^T(\vec{r}-\vec{R}_T) = Ck_0^{(1)}(\varkappa|\vec{r}-\vec{R}_T|) = C\frac{e^{-\varkappa|\vec{r}-\vec{R}_T|}}{\varkappa|\vec{r}-\vec{R}_T|} \tag{42}$$

where $|\vec{r}-\vec{R}_T| \neq 0$ since the solution is obtained in the vacuum at position $\vec{r}$ and C is a normalization constant.
Inserting this ansatz for the tip wave function into the expression for the matrix tunneling element (Eq. 35) yields

$$M_{\mu\nu}(\vec{R}_T) = -\frac{C\hbar^2}{2m}\int_{\Omega_T}\left[k_0^{(1)}(\varkappa|\vec{r}-\vec{R}_T|)\Delta\psi_\mu^S(\vec{r}-\vec{R}_T) - \psi_\mu^S(\vec{r}-\vec{R}_T)\Delta k_0^{(1)}(\varkappa|\vec{r}-\vec{R}_T|)\right]d\vec{r}. \tag{43}$$

Since the sample potential vanishes in the body of the tip we can apply the vacuum Schrödinger equation to the first term of the integrand. The second term has a singularity at $\vec{r} = \vec{R}$ and can be simplified recalling the relation between the modified Bessel function $k_0^{(1)}$ and the Green function of the vacuum Schrödinger equation:

$$\Delta G(\vec{r}-\vec{r}') = -4\pi\delta(\vec{r}-\vec{r}'). \tag{44}$$

Since $G(\vec{r}-\vec{r}') = \varkappa k_0^{(1)}(\varkappa|\vec{r}-\vec{r}'|)$, we rewrite $\Delta k_0^{(1)}$ in Eq. 43 as $(\varkappa^2 k_0^{(1)} - 4\pi\delta/\varkappa)$ and by that the tunneling matrix element for the case of a s-wave function simplifies in the TH model to

$$M_{\mu\nu}(\vec{R}_T) = -\frac{2\pi C\hbar^2}{\varkappa m}\psi_\mu^S(\vec{R}_T) \tag{45}$$

This is the central result of the TH model of STM although the original derivation is a bit longer. If the tip state is spherically symmetric around a point $\vec{R}_T$, effectively, it is equivalent to a geometrical point at $\vec{R}_T$. Hence, the tunneling matrix element is directly proportional to the value of the sample wave function at the position of the apex atom. Now we are able to calculate the tunneling current following Bardeen's formulation at low temperature:

$$I(\vec{R}_T, V) = \frac{16\pi^3 C^2\hbar^3 e}{\varkappa^2 m^2}n^T\int_0^{eV} d\epsilon n^S(\vec{R}_T, E_F^S+\epsilon) \tag{46}$$

where n^T is a constant in the TH model since the wave function of the tip is of s-type and $n^S(\vec{R}_T,\epsilon) = \sum_\mu \delta(E_\mu^S-\epsilon)|\psi_\mu^S(\vec{R}_T)|^2$. Eq. 46 expresses that the integral includes all states of the sample at the tip location between the Fermi energy and the Fermi energy shifted by the applied bias voltage: the tunneling current is proportional to the integrated local density of states (ILDOS) of the sample.

For a small bias voltage, the previous equation simplifies further to

$$I(\vec{R}_T, V) = \frac{16\pi^3 C^2 \hbar^3 e^2}{\varkappa^2 m^2} V n^T n^S(\vec{R}_T, E_F^S). \tag{47}$$

For the differential conductivity we obtain for finite V

$$\frac{dI(\vec{R}_T, V)}{dV} = \frac{16\pi^3 C^2 \hbar^3 e}{\varkappa^2 m^2} n^T n^S(\vec{R}_T, E_F^S + eV). \tag{48}$$

Thus the tunneling current and the differential conductivity are proportional to the density of states $n^S(\vec{R}_T, E_F^S + eV)$ in the vacuum. The last three equations are the most used ones in the interpretation, simulation and prediction of STM-images for realistic systems. We note that an absolute value of the current within this scheme cannot be computed since the constant C is unknown. From the TH model, one can simulate the important STM modes mentioned in Section 2. For example, in the constant-current mode (I = const.), where topographic images are obtained, we use $I \sim eVn^S(\vec{R}_T, E_F^S)$ valid for $eV << \Phi$. Hence, the task is simply to look for $n^S(\vec{R}_T, E_F^S) = \text{const.}$ In the spectroscopic mode, dI/dV is computed from $\sim n^S(\vec{R}_T, E_F^S + eV)$, in other words the calculation of the spectroscopic images obtained with STM boils down to the computation of the sample's DOS in vacuum.
The basic assumption of this extremely simple result is that, except for the s-wave tip wave function, all other tip wave functions can be neglected. Therefore it is often called the s-wave tip model. It is important to know under which condition the s-wave-only assumption is valid. The TH model is highly valuable in the interpretation of STM images. It represents an approximation with which the complicated problem of tip electronic states can be avoided.
What if the STM-tip is more complicated than what is assumed in the TH-model? For example, other orbitals, than the s, can be characterizing the apex atom. Chen [25, 26] proposed an elegant method, discussed in the upcoming subsection, that extends the TH-model.

Chen's expansion of the Tersoff-Hamann model

The problems which arise from the Tersoff-Hamann model can be overcome by expanding the model using generalized wave functions for the tip. Such an expansion was introduced by Chen [25, 26] who considered the general solutions of the vacuum Schrödinger equation (Eq. 41)

$$\psi^T(\vec{r} - \vec{R}_T) = \sum_{l,m} C_{lm} k_l^{(1)}(\varkappa|\vec{r} - \vec{R}_T|) Y_{lm}(\widehat{|\vec{r} - \vec{R}_T|}) \tag{49}$$

where Y_{lm} are the spherical harmonics and $k_1^{(1)}$ are modified spherical Bessel functions of the second kind while C_{lm} is a renormalization coefficient. For the case $l = 0$ we recover the TH model as detailed in the previous subsection. Using the property of the Bessel function:

$$k_l^{(1)}(u) = (-1)^l u^l \left(\frac{1}{u}\frac{d}{du}\right)^l k_0^{(1)}(u) \tag{50}$$

one can evaluate the contribution of a tip-orbital l to the tunneling current just by proceeding to the l^{th} derivative with respect to the argument of modified Bessel function with $l = 0$. For

example, if the tip is described by a p_z-type orbital we have

$$\begin{aligned}\psi^T_{p_z}(\vec{(}r) - \vec{R}_T) &= C_{p_z} k_1^{(1)}(\varkappa|\vec{r} - \vec{R}|) Y_{10}(\widehat{|\vec{r} - \vec{R}_T|}) \\ &= C_{p_z} \frac{d}{dR} k_0^{(1)}(\varkappa|\vec{r} - \vec{R}_T|) \frac{\partial R}{\partial z} \\ &= \frac{C_{p_z}}{C_s} \frac{\partial}{\partial z} \psi^T_s(\vec{r} - \vec{R}_T).\end{aligned} \tag{51}$$

Inserting this wave function in the Bardeen's formula for the tunneling matrix elements in the TH model, we obtain for the p_z orbital

$$M_{\mu\nu} = -\frac{2\pi C_{p_z}\hbar^2}{\varkappa m} \frac{\partial}{\partial z} \psi^S_\mu(\vec{R}_T). \tag{52}$$

This means, that the matrix element is proportional to the derivative of the sample wave function with respect to Z at the position of the tip, if the tip is described by a p_z-type orbital. In this way the matrix element can be derived also for higher order orbitals, which is known as the derivative rule of Chen [25, 26].
Using p- or d-type orbitals for the tip, the experimentally observed corrugation amplitudes of densely packed metal surfaces can be explained. In Table 4.2 the tunneling matrix elements are given for different orbitals. With the help of this rule, Chen could explain the high corrugation amplitude observed in some systems.
The extension of Chen provides an explanation of the high corrugation amplitudes measured on close-packed metal surfaces contradictory to the low corrugation amplitudes due to their local density of states. This is the case, for example, for p_z and d_{z^2} orbitals since they possess charge density stretching out further from the tip apex into the vacuum than that of an s-wave and they act similar to an s-wave at a reduced distance from the sample surface. This affects the tunneling current quite strongly and by that the images obtained experimentally. Interestingly, orbitals like d_{xy} and $d_{xz,yz}$ are expected to produce a large tunneling current not with the tip apex atom located on top of a surface atom but rather at a hollow site of the surface. Due to their particular charge density distribution a large overlap with sample orbitals occurs in this configuration.

4.3 Different STM modes

Now that we know how to compute the tunneling current, it is interesting to connect this quantity to the different STM standard modes. For a given bias, the physical quantity measured by the STM is the tunneling current, which is a function of the lateral coordinates (x, y) and the z-coordinate: $I = I(x, y, z)$. If z is perpendicular to a nearly flat surface, the tunneling current can be decomposed into a constant (that is independent of (x, y)) and a small variable component that represents the features or corrugation of the surface,

$$I(x, y, z) = I_0(z) + \Delta I(x, y, z), \tag{53}$$

with the condition $|\Delta I(x, y, z)| << |I_0(z)|$.
The constant-current topographic image can be derived from the current images by making the ansatz: $z(x, y) = z_0 + \Delta z(x, y)$ and substituting it into the previous equation by proceeding to a Taylor expansion we find

$$I = I_0(z_0) + \left(\frac{dI_0(z)}{dz}\right) \Delta z(x, y) + \Delta I(z, y, z). \tag{54}$$

Orbital of the tip	Matrix element M
s	$\frac{2\pi C\hbar^2}{\varkappa m}\psi^S(\vec{R}_T)$
p_x	$\frac{2\pi C\hbar^2}{\varkappa m}\frac{\partial}{\partial x}\psi^S(\vec{R}_T)$
p_y	$\frac{2\pi C\hbar^2}{\varkappa m}\frac{\partial}{\partial y}\psi^S(\vec{R}_T)$
p_z	$\frac{2\pi C\hbar^2}{\varkappa m}\frac{\partial}{\partial z}\psi^S(\vec{R}_T)$
d_{zx}	$\frac{2\pi C\hbar^2}{\varkappa m}\frac{\partial^2}{\partial z\partial x}\psi^S(\vec{R}_T)$
d_{zy}	$\frac{2\pi C\hbar^2}{\varkappa m}\frac{\partial^2}{\partial z\partial y}\psi^S(\vec{R}_T)$
d_{xy}	$\frac{2\pi C\hbar^2}{\varkappa m}\frac{\partial^2}{\partial x\partial y}\psi^S(\vec{R}_T)$
$d_{z^2-\frac{1}{3}r^2}$	$\frac{2\pi C\hbar^2}{\varkappa m}\left(\frac{\partial^2}{\partial z^2}\psi^S(\vec{R}_T)-\frac{1}{3}\varkappa^2\psi^S(\vec{R}_T)\right)$
$d_{x^2-y^2}$	$\frac{2\pi C\hbar^2}{\varkappa m}\left(\frac{\partial^2}{\partial x^2}\psi^S(\vec{R}_T)-\frac{\partial^2}{\partial y^2}\psi^S(\vec{R}_T)\right)$

Table 2: *Tunneling matrix elements as formulated by Chen [22] using the derivative rule. Note that the constant C depends on the orbital-type of the tip involved in the tunneling process.*

Owing to the smallness of ΔI, its variation can be neglected. The topographic image is therefore defined by the condition of constant current, i.e. $I = I_0(z_0)$, thus

$$\Delta z(x,y) = -\frac{\Delta I(x,y)}{dI_0(z)/dz}. \tag{55}$$

It is interesting to note that experimentally when scanning the sample surface with the tip there are two different modes of operation, the constant-height and the constant-current mode. In constant-height mode, the vertical position z of the tip is held constant while scanning and the resulting tunnel current between tip and sample is measured. In constant-current mode a feedback loop provides a constant tunnel current between tip and sample at every position (x, y). This means that the z-position of the tip has to be adjusted during scanning which is done by applying an appropriate voltage V_z to the z-piezo of the tube scanner. One distinguishes between these two extreme modes of operation even though neither of them can be realized experimentally and one can only approximate one or the other by choosing the appropriate parameters for the feedback loop gain and the scan speed.
The differential tunneling conductivity, dI/dV, is also a frequently used image parameter. Experimentally the tunneling spectrum at each point (x, y) can be obtained by interrupting the feedback circuit, applying a voltage ramp, then acquiring the tunneling current. The differential tunneling conductivity can be obtained by numerically differentiating the acquired tunneling current data.

4.4 Spin-polarization and tunneling: SP-STM and TAMR-STM

If the tunnel current flows between two magnetic electrodes, an additional information will be contained in the tunneling current, namely the information on the magnetic properties of the electrodes. Thus, if the STM tip is spin-polarized, in other words, the DOS for the majority-spin (↑) channel is different from the DOS for the minority-spin (↓) channel, access to the local spin-polarization of the probed substrate is possible (Figs. 5(a) and (b)). This is the concept of spin-polarized STM (SP-STM). By assuming that the electron spin is conserved during the tunneling process, the ↑-electrons from the tip can only tunnel into unoccupied ↑-states in the sample; the same for the ↓-electrons (Fig. 5(b)). When the magnetization directions of the two electrodes are in parallel alignment the tunnel current is different compared to the antiparallel alignment.
In fact Jullière [48] first discovered this spin valve effect, or tunneling magnetic resistance (TMR), in planar Fe-Ge-Co tunnel junctions effect which showed a decreased conductance in the case of non-parallel alignment of the electrodes magnetization compared to the parallel case. In a theoretical work, Slonczewski extended the model of tunneling in one dimension considering spin-polarized electrodes [49]. While the experiments of Jullière had to be realized at very low temperatures, the TMR effect at room temperature was achieved in the 90's by Moodera [50] and Miyazaki [51]. This allows for the application of the TMR effect in read heads of modern hard disk drives. Furthermore, the discovery of the TMR gave rise to the development of the magneto-resistive random-access memory (MRAM) – a non-volatile random-access memory technology.
Due to spin-orbit coupling (SOC), the resistance can become anisotropic, i.e., it depends on the magnetization direction of the tunnel junction with respect to the crystallographic axes as sketched in Fig. 5(c). For the observation of this effect the tunneling junction needs only a single magnetic electrode separated from a nonmagnetic electrode by an insulating layer. In

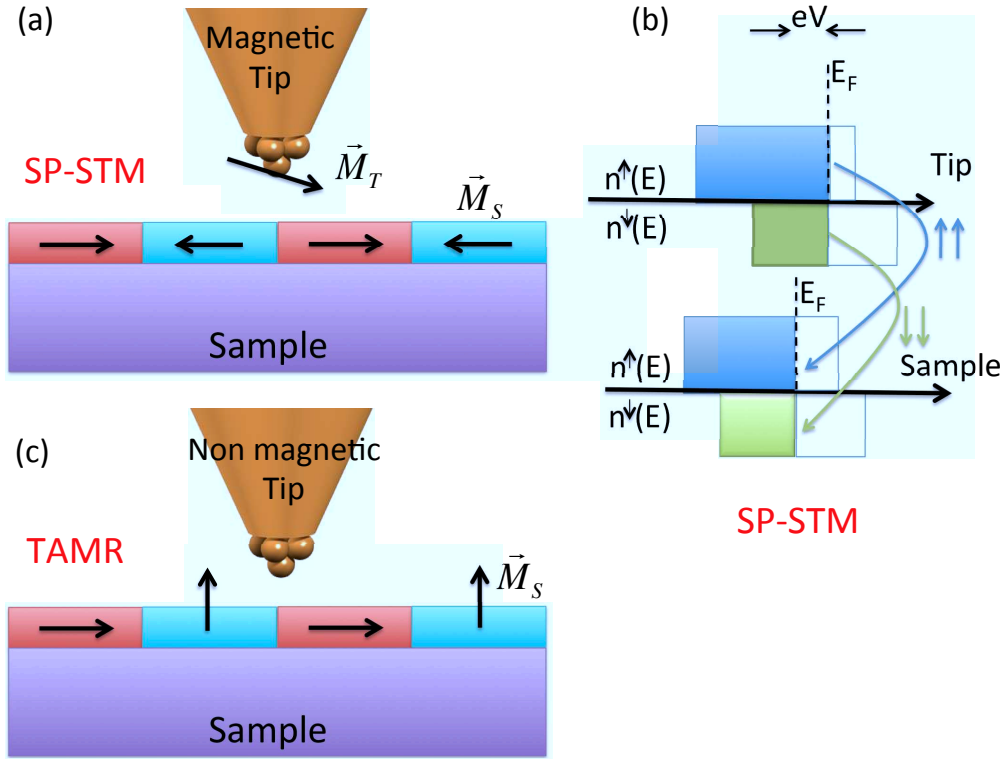


Fig. 5: *Spin-polarized tunneling with SP-STM shown schematically in (a) and (b). In (b), the spin-polarized DOS of the tip and sample are depicted. With spin conservation, electrons from the tip with spin ↑ can tunnel into unoccupied states of the sample with the same spin character. The same tunneling process occurs with spin ↓. In (c) a sketch of TAMR is shown. The tunnel junction comprises a non-magnetic tip and a magnetic sample. Because of spin-orbit coupling the tunneling current can be sensitive to the magnetization's orientation of the sample.*

fact depending on the magnetization direction, the electronic structure of the magnetic electrode changes, thus the tunneling current between the two electrodes exhibits differences for a film that is magnetized either out-of-plane or in-plane [52, 53]. This effect has been named tunneling anisotropic magneto-resistance (TAMR) [53] which is an extension of the known bulk AMR that does not involve tunneling. Besides its implications in spintronics, it is appealing to use the TAMR concept in STM since even without a spin-polarized tip, magnetic information can be grasped from the tunneling current if the sample is characterized by a non-negligible SOC. The TAMR has first been observed in STM measurements of a double-layer film of Fe on the W(110) surface [52] and was recently applied on adatoms deposited on magnetic substrates [54, 55].
Since more than 10 years, SP-STM is a well-established technique which can be used to investigate the magnetic ground state of nanostructures down to the atomic scale [9]. In the past years, the technique has been extended to study also dynamics of magnetic systems like spin-flip processes [15, 16, 17, 18] or magnon excitation [57]. Furthermore, very recently SP-STM has been used to probe spin relaxations of single atoms on the time scale of nanoseconds [58, 59] to few minutes [60].
In the following two subsections, the basic theoretical concepts behind SP-STM and TAMR

within STM are presented.

Bardeen's formalism for SP-STM

The Bardeen tunneling theory can be extended to include the spin dependence. Instead of using a single-component wave function, two components, i.e. a spinor, are necessary to describe a state of an electron with spin. For example, for the sample wave function

$$\psi^S_{\mu\sigma} = \sum_{\sigma'=\uparrow,\downarrow} \psi^s_{\mu\sigma\sigma'}\chi_{\sigma'}, \text{ with } \chi_\uparrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \chi_\downarrow = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{56}$$

One can follow the same procedure as done previously in Section 4.1 for the non spin-polarized case by considering the time-dependent Pauli equation of the combined system

$$i\hbar\frac{\partial\psi}{\partial t} = \left[-\frac{\hbar^2}{2m}\nabla^2 + \hat{U}_T + \hat{U}_S\right]\psi. \tag{57}$$

Note that the wave function ψ is now a two-component spinor and the potential functions are now two-by-two matrices,

$$\hat{U}_T = \begin{pmatrix} U_{T\uparrow\uparrow} & U_{T\uparrow\downarrow} \\ U_{T\downarrow\uparrow} & U_{T\downarrow\downarrow} \end{pmatrix}, \hat{U}_S = \begin{pmatrix} U_{S\uparrow\uparrow} & U_{S\uparrow\downarrow} \\ U_{S\downarrow\uparrow} & U_{S\downarrow\downarrow} \end{pmatrix}, \tag{58}$$

We follow the treatment of Wortmann *et al*. [61] to take the spin-polarization direction of one of the electrodes, say, electrode T, as the reference (global spin frame of reference). In other words, we assume that the Hamiltonian of the tip is diagonal with respect to spin,

$$\hat{U}_T = \begin{pmatrix} U_{T\uparrow\uparrow} & 0 \\ 0 & U_{T\downarrow\downarrow} \end{pmatrix}. \tag{59}$$

Therefore, the two components of the wave functions of the tip-only system can be treated separately to satisfy the following equation (considering the tip only),

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + U_{T\sigma\sigma}\right]\psi^T_{\nu\sigma\sigma}(\vec{r}) = E_{\nu\sigma\sigma}\psi^T_{\nu\sigma\sigma}(\vec{r}), \tag{60}$$

where σ denotes the spin component $\uparrow$ or $\downarrow$.
However, in the reference frame of the tip, the state of electrode S is, in general, not diagonalized with respect to spin. This is evidently true for non-collinear systems since no quantization axis exists which allows a state to be written in terms of pure spin-up or spin-down character, but even for collinear samples the states will be spin mixed if the quantization axis of the sample and the one of the tip are not aligned in parallel. In general, the spinor of the sample-only system satisfies the Pauli equation,

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + \begin{pmatrix} U_{S\uparrow\uparrow} & U_{S\uparrow\downarrow} \\ U_{S\downarrow\uparrow} & U_{S\downarrow\downarrow} \end{pmatrix}\right]\begin{pmatrix} \psi^S_{\mu\uparrow\sigma} \\ \psi^S_{\mu\downarrow\sigma} \end{pmatrix} = \hat{E}_\mu \begin{pmatrix} \psi^S_{\mu\uparrow\sigma} \\ \psi^S_{\mu\downarrow\sigma} \end{pmatrix}. \tag{61}$$

It is useful to express the components of spinor describing the sample's wave function in the spin-frame of reference of the tip within the local spin frame of reference related to the sample.

In this local spin-frame of reference $\hat{U}_S$ is diagonal in spin-space. This can be achieved using the rotation matrix $\hat{U}$.

$$|\psi_\sigma^S\rangle = \hat{U}(\theta)|\psi_\sigma^{S,loc}\rangle, \text{ with } \hat{U}(\theta) = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \tag{62}$$

where loc stands for local spin-frame of reference of the sample and θ defines the angle between the magnetization of the individual atom on the sample and the magnetization of the tip.
Following the same procedure as done in the non spin-polarized case, we find in first-order perturbation theory that the spin-dependent tunneling matrix elements are given by

$$M_{\mu\nu}^{\sigma\sigma'} = \left\langle \psi_{\mu\sigma'}^T \left| \hat{U}_T \hat{U}(\theta) \right| \psi_{\mu\sigma}^{S,loc} \right\rangle \tag{63}$$

and

$$\begin{pmatrix} M_{\mu\nu}^{\uparrow\uparrow} & M_{\mu\nu}^{\uparrow\downarrow} \\ M_{\mu\nu}^{\downarrow\uparrow} & M_{\mu\nu}^{\downarrow\downarrow} \end{pmatrix} = \begin{pmatrix} \langle\psi_{\mu\uparrow}^T|U_{T\uparrow\uparrow}|\psi_{\mu\uparrow}^{S,loc}\rangle\cos(\theta/2) & -\langle\psi_{\mu\uparrow}^T|U_{T\uparrow\uparrow}|\psi_{\mu\downarrow}^{S,loc}\rangle\sin(\theta/2) \\ \langle\psi_{\mu\downarrow}^T|U_{T\downarrow\downarrow}|\psi_{\mu\uparrow}^{S,loc}\rangle\sin(\theta/2) & \langle\psi_{\mu\downarrow}^T|U_{T\downarrow\downarrow}|\psi_{\mu\downarrow}^{S,loc}\rangle\cos(\theta/2) \end{pmatrix} \tag{64}$$

$$= -\frac{2\pi\hbar^2}{m} \begin{pmatrix} \frac{C_\uparrow}{\varkappa_\uparrow}\psi_{\mu\uparrow}^{S,loc}(\vec{R}_T)\cos(\theta/2) & -\frac{C_\uparrow}{\varkappa_\uparrow}\psi_{\mu\downarrow}^{S,loc}(\vec{R}_T)\sin(\theta/2) \\ \frac{C_\downarrow}{\varkappa_\downarrow}\psi_{\mu\uparrow}^{S,loc}(\vec{R}_T)\sin(\theta/2) & \frac{C_\downarrow}{\varkappa_\downarrow}\psi_{\mu\downarrow}^{S,loc}(\vec{R}_T)\cos(\theta/2) \end{pmatrix} \tag{65}$$

In the last equation we followed the TH model to extract the tunneling matrix elements by replacing the wave function at the tip apex atom by a spherically symmetric s-wave. Also the Chen's rule for arbitrary orbitals can be followed in the spin-polarized case. In the following we assume that the spin-up and spin-down s-wave states can be characterized by the same decay constant $\varkappa$ and the same normalization coefficient C, i.e. $\varkappa_\sigma = \varkappa$ and $C_\sigma = C$.
The tunneling current in the spin-polarized case becomes

$$\begin{aligned} I(\theta,V) = {} & \frac{2\pi e}{\hbar}\sum_{\sigma\sigma'}\int_0^{eV} d\epsilon \left[f(E_F^T - eV + \epsilon) - f(E_F^S + \epsilon)\right] \\ & \times n_{\sigma'}^T(E_F^T - eV + \epsilon) n_\sigma^S(E_F^S + \epsilon) \left|M_{\sigma\sigma\prime}(\theta, E_F^S + \epsilon, E_F^T - eV + \epsilon)\right|^2 \end{aligned} \tag{66}$$

In the TH model, the tunneling current simplifies to

$$\begin{aligned} I(\vec{R}_T,\theta,V) = {} & \frac{16\pi^3 C^2\hbar^3 e}{\varkappa^2 m^2}\int_0^{eV} d\epsilon \left[f(E_F^T - eV + \epsilon) - f(E_F^S + \epsilon)\right] \\ & \times \left(n^T(E_F^T - eV + \epsilon) n^S(\vec{R}_T, E_F^S + \epsilon) + \vec{m}^T(E_F^T - eV + \epsilon)\cdot\vec{m}^S(\vec{R}_T, E_F^S + \epsilon)\right) \end{aligned} \tag{67}$$

where n is the total charge density of states ($n = n_\uparrow + n_\downarrow$) of the tip or sample while $\vec{m}$ is the magnetization vector, i.e. the corresponding "magnetic" density of states. A further approximations can be made by considering the tip to be characterized by an s-wave function. That allows to consider n^T to be a constant:

$$I(\vec{R}_T,\theta,V) = \frac{16\pi^3 C^2\hbar^3 e}{\varkappa^2 m^2}(\underbrace{n^T N^S(\vec{R}_T,V)}_{\text{nonspin-polarized}} + \underbrace{\vec{m}^T\cdot\vec{M}^S(\vec{R}_T,V)}_{\text{spin-polarized}}), \tag{68}$$

where N^s and M^s are respectively the energy integrated charge-density and magnetization of the sample's atom at the tip location $\vec{R}_T$. The previous equation, widely used to interpret SP-STM, shows that the tunneling current can be decomposed into a non spin-polarized and spin-polarized contributions. In case of a non spin-polarized STM experiment, i.e., using either a nonmagnetic tip or sample, the second term vanishes and the current reduces to the classical result of the TH model. Furthermore, depending on the angle between the magnetization vectors of tip and sample, the current will change.
For completeness we give the corresponding differential conductivity in the spin-polarized case

$$\frac{dI}{dV}(\vec{R}_T,\theta,V) \propto (\underbrace{n^T n^S(\vec{R}_T,E_F+eV)}_{\mathrm{nonspin-polarized}} + \underbrace{\vec{m}^T\cdot\vec{m}^S(\vec{R}_T,E_F+eV)}_{\mathrm{spin-polarized}}). \tag{69}$$

Tunneling anisotropic magneto-resistance (TAMR)

We derived previously the tunneling current and the differential conductivity in case of a spin-polarized tip. If the tip is non spin-polarized, the current depends solely on the non spin-polarized part. Assume that the sample magnetization is oriented along the out-of-plane direction, and that one applies a magnetic field to reorient the magnetization to be in-plane. In some cases, for instance if SOC is present, this reorientation could affect the electronic structure [56]. In other words, this means that n^S then exhibits a dependence on the magnetization's orientation.
This can be noticed by considering the SOC potential

$$\hat{V}_{SOC} = \zeta\vec{L}\cdot\hat{\vec{S}} = \zeta\begin{pmatrix} V_{SOC}^{\uparrow\uparrow} & V_{SOC}^{\uparrow\downarrow} \\ V_{SOC}^{\downarrow\uparrow} & V_{SOC}^{\downarrow\downarrow}\end{pmatrix} \tag{70}$$

where ζ is the SOC strength, $\vec{L}$ and $\vec{S}$ are the orbital and angular momenta. The Hamiltonian without SOC, H^0, is spin-diagonal and the unperturbed Bloch eigenfunctions are $(\psi^{(0)}_{\vec{k}\mu\uparrow},0)^T$ and $(0,\psi^{(0)}_{\vec{k}\mu\downarrow})^T$. The Schrödinger equation for the perturbed wave function then reads

$$\begin{pmatrix} H^{0\uparrow}+V_{SOC}^{\uparrow\uparrow}-E & V_{SOC}^{\uparrow\downarrow} \\ V_{SOC}^{\downarrow\uparrow} & H^{0\downarrow}+V_{SOC}^{\downarrow\downarrow}-E\end{pmatrix}\begin{pmatrix}\psi_{\vec{k}\mu\sigma\uparrow}\\ \psi_{\vec{k}\mu\sigma\downarrow}\end{pmatrix} = 0. \tag{71}$$

The potential terms $V_{SOC}\uparrow\downarrow$ and $V_{SOC}\downarrow\uparrow$ are responsible for flipping the spin. Solving this equation for the minority-spin channel, i.e. $\sigma=\downarrow$, leads for the two components of the spinor in first-order perturbation theory:

$$\psi^{(1)}_{\vec{k}\mu\downarrow\uparrow} = \sum_\nu \frac{\langle\psi^{(0)}_{\vec{k}\nu\downarrow}|V_{SOC}^{\downarrow\uparrow}|\psi^{(0)}_{\vec{k}\mu\uparrow}\rangle}{E^{0\uparrow}_{\vec{k}\mu}-E^{0\downarrow}_{\vec{k}\nu}}\psi^{(0)}_{\vec{k}\nu\downarrow} \tag{72}$$

and

$$\psi^{(1)}_{\vec{k}\mu\downarrow\downarrow} = \psi^{(0)}_{\vec{k}\mu\downarrow} + \sum_{\nu\neq\mu} \frac{\langle\psi^{(0)}_{\vec{k}\nu\downarrow}|V_{SOC}^{\downarrow\downarrow}|\psi^{(0)}_{\vec{k}\mu\downarrow}\rangle}{E^{0\downarrow}_{\vec{k}\mu}-E^{0\downarrow}_{\vec{k}\nu}}\psi^{(0)}_{\vec{k}\nu\downarrow} \tag{73}$$

where the index (1) stands for the first-order solution and $E^{0\uparrow}_{\vec{k}\mu}$ the eigenenergy of the state $\psi^0_{\vec{k}\mu\uparrow}$. Since the DOS is related to $|\psi^{(1)}|^2$, it is expected that the DOS will change because of SOC in

a quadratic fashion: $\sim \sum_{\sigma\sigma'}(V_{SOC}^{\sigma\sigma'})^2$ but the denominator of the previous equations will play an important role. Indeed if the weight of the states close to those $\vec{k}$ where the unperturbed spin-dependent bands cross, i.e. $E_{\vec{k}\mu}^{0\sigma} = E_{\vec{k}\mu}^{0\sigma'}$ the denominator becomes small and the bands strongly couple and the modification of the charge can be important.

To illustrate the dependence of the DOS on the SOC and rotation angle of the magnetization, we introduce a simple toy model initially proposed by Néel *et al.* [54, 62] that shows how the total DOS, and by that the tunneling current and differential conductivity, can be affected by the magnetization orientation.

Assume that in the energy window of interest, we have two states, say d_{z^2} and d_{zx} states, in the minority spin-channel originally located at energies ϵ_1 and ϵ_2. For simplification, we consider full spin-polarization, meaning that no orbitals are present in the majority spin channel. The hybridization with the background is described via Γ_1 and Γ_2. The interaction between the two terms is provided via a hopping term t which in our case is created by SOC. The form of t is inferred from the work of Abate and Asdente [63] on bulk Fe who used a tight-binding formalism to evaluate the matrix elements $|\langle d_m|V_{SOC}|d_{m'}\rangle|$.

The matrix element connecting our orbitals are given by

$$|\langle d_{zx}|V_{SOC}|d_{z^2}\rangle| = \frac{1}{2}\sqrt{3}\zeta \sin\theta \sin\phi \tag{74}$$

for the minority spin-channel. θ and ϕ are the Euler angles defining the orientation of the magnetization.

We calculate the density state of this two-orbitals model by evaluating the Green function $G = (1-H)^{-1}$ from which the imaginary part is extracted ($n(E) = -\frac{1}{\pi}\sum_{m=1}^{2} \Im G_{mm}(E)$). We find

$$\hat{G} = \frac{1}{(E-\epsilon_1-i\Gamma_1)(E-\epsilon_2-i\Gamma_2)-t^2}\begin{pmatrix} E-\epsilon_1-i\Gamma_1 & t \\ t & E-\epsilon_2-i\Gamma_2 \end{pmatrix}, \tag{75}$$

where one notices that the diagonal elements of the Green functions, G_{mm} are proportional to t^2, i.e., to V_{soc}^2 which bears the angular dependence. Thus this toy model indicates, that in the presence of SOC, the local DOS will depend on the orientation angle of the magnetization vector. The results obtained with ζ =50 meV (approximate value for $3d$ transition elements) are depicted in Fig. 6(a) where the corresponding DOS is shown for two different orientations, out-of-plane ($\theta = 0^o$) and in-plane ($\theta = 90^o, \phi = 0^o$). One notices that changes occur for both orbitals upon rotation. From our different formulas giving the tunneling current it is obvious that since the charge of the sample gets modified, the current magnitude will be affected by the rotation of the magnetization vector. In Fig. 6(b) is presented the corresponding TAMR signal, given by $\frac{n(\theta=0)-n(\theta=90)}{n(\theta=0)}$. A value of 20% is found at the position of the d-resonances where the large change in the DOS is observed. Experimentally, it is thus worthwhile to probe the sample at those energies where the largest TAMR effect is observed. It is interesting to check the angle dependence of the DOS as shown in Fig. 6(c). Interestingly, the d_z^2 contribution to the DOS, thus the tunneling current, versus the rotation angle follows a $\sin^2$ behavior in contrast to the cosine behavior of the spin-polarized part of the tunneling current in the SP-STM geometry when SOC is not included.

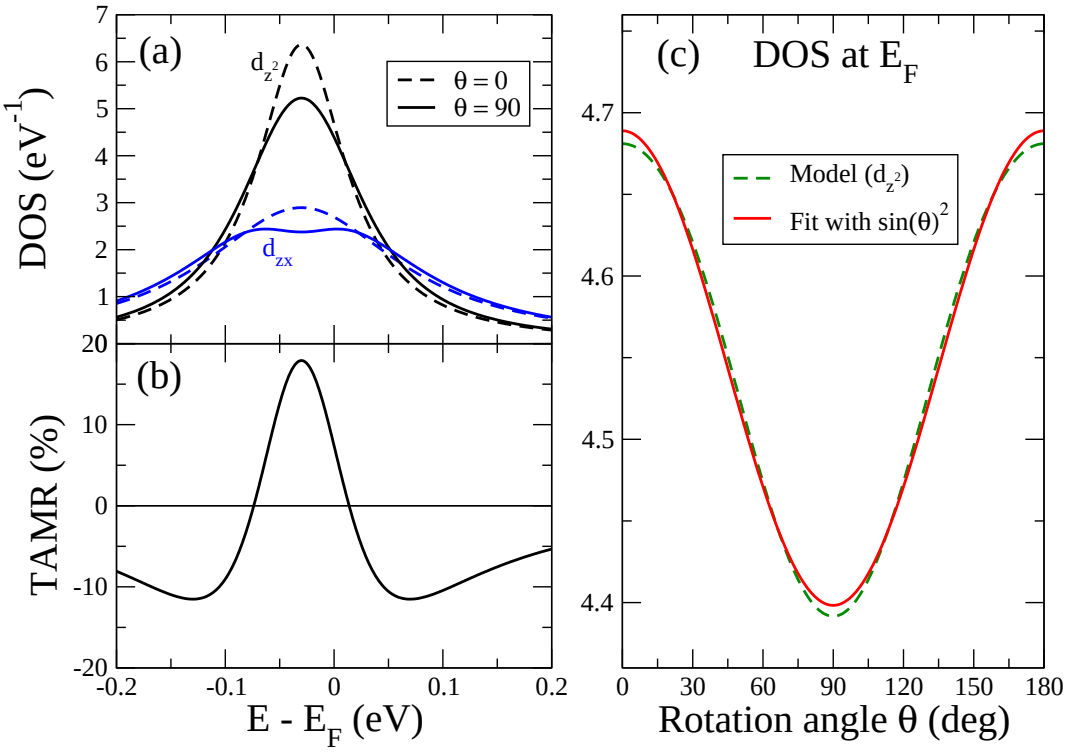


Fig. 6: *(a) DOS obtained from a simple model for two d orbitals interaction via SOC (see Refs.[54, 62]). Dashed and full lines refer to the DOS calculated when the magnetization is out-of-plane and in-plane. The corresponding TAMR signal is plotted in (b) while the variation of the d_{z^2}-DOS at E_F with respect to the rotation angle θ is shown in (c).*

4.5 Crystal surfaces: $k_{||}$-Selection in STM

The DOS in vacuum above the substrate can be computed with ab-initio method, but for bulk systems before getting the DOS in real space a summation over k–points in reciprocal space has to be performed within the Brillouin zone. It is instructive to realize that not all k-points contribute equally to the vacuum's DOS. Indeed on a surface probed by STM, because of symmetry reduction only the parallel component, $\vec{k}_{||}$, of the the three dimensional Bloch vector $\vec{k} = (k_\perp, \vec{k_{||}})$ remains as a good quantum number. Thus in vacuum the wave function, $\psi_{\vec{k}_{||}\mu}$, describing the surface characterized on the basis of Bloch theorem by a band index μ and a wave-vector $\vec{k}_{||}$ of the two-dimensional (2D) Brillouin zone is expanded into basis functions:

$$\psi_{\vec{k}_{||}\mu}(\vec{r}_{||}, z) = \sum_{\alpha} c^n_{\vec{k}_{||}\mu} d^n_{\vec{k}_{||}\mu}(z) \exp\left[i(\vec{k}_{||} + \vec{G}^n_{||})\vec{r}_{||}\right] \tag{76}$$

which are 2D plane waves parallel to the surface. $\vec{G}^n_{||}$ denotes the reciprocal lattice vectors parallel to the surface while $d^\alpha_{\vec{k}_{||}}$ are basis functions describing the decay of the substrate's states into vacuum and can be obtained by solving the one-dimensional Schrödinger equation

in vacuum

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dz^2}+U(z)-\epsilon_\mu+\frac{\hbar^2}{2m}(\vec{k}_{||}+\vec{G}_{||}^{\alpha})^2\right)d^{\alpha}_{\vec{k}_{||}}(z)=0, \tag{77}$$

for a given vacuum potential $U(z)$ and reference energy ϵ_μ.
In general, quantities possessing the crystal symmetry of the lattice can be expanded into a set of symmetrized functions. The local density of states (LDOS)

$$\begin{aligned} n(\vec{r}_{||},z;\epsilon) &= \sum_{\vec{k}_{||}\mu}\delta(\epsilon-\epsilon_{\vec{k}_{||}\mu})|\psi_{\vec{k}_{||}\mu}(\vec{r}_{||},z)|^2 \\ &= \sum_{\vec{k}_{||}\mu}\delta(\epsilon-\epsilon_{\vec{k}_{||}\mu})\sum_{\beta} n^{\beta}_{\vec{k}_{||}\mu}(z)\exp\left(i\vec{G}_{||}^{\beta}\vec{r}_{||}\right) \end{aligned} \tag{78}$$

where $\exp i\vec{G}_{||}^{\beta}\vec{r}_{||}$ are called star coefficients and

$$n^{\beta}_{\vec{k}_{||}\mu}(z)=\sum_{\alpha\alpha'}c^{\alpha}_{\vec{k}_{||}\mu}c^{\alpha'*}_{\vec{k}_{||}\mu}d^{\alpha}_{\vec{k}_{||}\mu}(z)d^{\alpha'*}_{\vec{k}_{||}\mu}(z)\delta(\vec{G}_{||}^{\alpha}-\vec{G}_{||}^{\alpha'},\vec{G}_{||}^{\beta}) \tag{79}$$

d could be the s-orbital as proposed in TH model, or one can use arbitrary orbitals according to the derivative rule of Chen.
One could grasp the behavior of the LDOS in vacuum by considering for simplicity that $U(z)=0$. Thus the exact solution of the one dimensional Schrödinger equation in vacuum gives $d^{\alpha}_{\vec{k}_{||}}=\exp\left(-\varkappa^{\alpha}_{\vec{k}_{||}}z\right)$ where $z>0$. The decay constant

$$\varkappa^{\alpha}_{\vec{k}_{||}}=\sqrt{2m|\epsilon_\mu|/\hbar^2+(\vec{k}_{||}+\vec{G}_{||}^{\alpha})^2} \tag{80}$$

and the related LDOS

$$n^{\beta}_{\vec{k}_{||}\mu}(z)=\sum_{\alpha\alpha'}c^{\alpha}_{\vec{k}_{||}\mu}c^{\alpha'*}_{\vec{k}_{||}\mu}\exp\left[-(\varkappa^{\alpha}_{\vec{k}_{||}\mu}+\varkappa^{\alpha'}_{\vec{k}_{||}\mu})z\right]\delta(\vec{G}_{||}^{\alpha}-\vec{G}_{||}^{\alpha'},\vec{G}_{||}^{\beta}) \tag{81}$$

show obviously a strong dependence on $\vec{k}_{||}$ and on $\vec{G}_{||}^{\alpha}$.
The last equations demonstrate that the decay constant is the largest and thereby the LDOS is the smallest when contributions of $|\vec{k}_{||}+\vec{G}_{||}^{\alpha}|$ are significant. A decrease in $\varkappa$ of 0.1 Å^{-1} could lead to a reduction of the LDOS and of the tunneling current of the order of 50%.
Also the nature of the lattice can affect tunneling, as demonstrated in the following example. We show the first three vectors responding to the smallest "stars" (m=1, 2, 3) of reciprocal lattice vectors of a square lattice (see Fig. 7), which would represent the bcc(001) or fcc(001) surfaces. The corresponding reciprocal lattice vectors are $\vec{G}_{||}^{(1)}=(0,0)$, $\vec{G}_{||}^{(2)}=(1,0)$, and $\vec{G}_{||}^{(3)}=(1,1)$, expressed in units of $2\pi/a$, with a being the lattice constant and the star coefficients are:

$$\exp\left(i\vec{G}_{||}^{(1)}\vec{r}_{||}\right) = 1 \tag{82}$$

$$\exp\left(i\vec{G}_{||}^{(2)}\vec{r}_{||}\right) = \frac{1}{2}[\cos\left(\vec{G}_{||,1}\vec{r}_{||}\right)+\cos\left(\vec{G}_{||,2}\vec{r}_{||}\right)] \tag{83}$$

$$\exp\left(i\vec{G}_{||}^{(3)}\vec{r}_{||}\right) = \cos\left[(\vec{G}_{||,1}+\vec{G}_{||,2})\vec{r}_{||}\right]. \tag{84}$$

$\vec{G}_{||,1}=\frac{2\pi}{a}(1,0)$ and $\vec{G}_{||,2}=\frac{2\pi}{a}(0,1)$ are the two-dimensional reciprocal lattice vectors. In Fig. 7 the star functions are displayed together with the 2D unit cell for a checkerboard structure

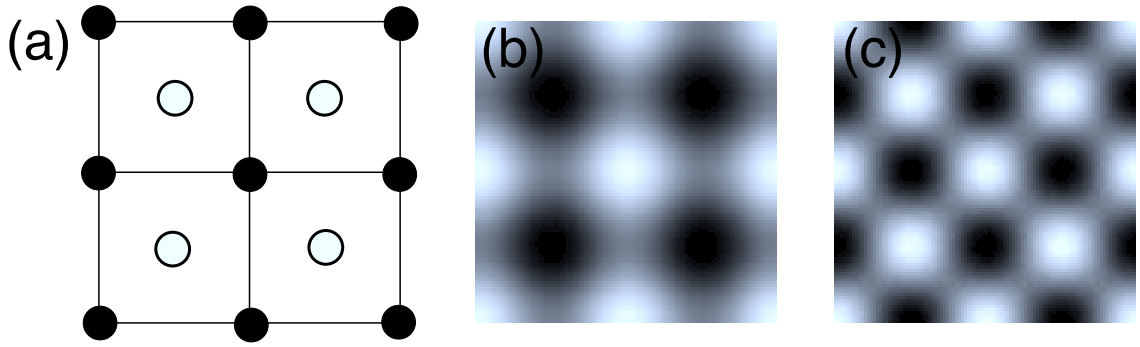


Fig. 7: *Star functions for a square lattice according to Ref. [38]. (a) unit cell of a square lattice with a two-atoms basis, i.e. representing a checkerboard structure. The first star function is a constant contrary to the second one which leads to a chemical sensitivity on the two atoms (b). The black atoms are displayed as protrusions and the white ones as depressions. Interestingly the third star function shows equally both atoms (c).*

with two different atom types. The first star function is a constant and represents the lateral constant part of the LDOS and thereby also of the tunneling current in the TH-model. Higher star coefficients can contribute non-trivially to the final STM image. The second coefficient for example, allows to distinguish between the two kinds of atoms while the third coefficient does not, i.e., chemical sensitivity is probed only by the second star coefficient.
Any magnetic superstructure lowers the translational symmetry. Therefore, smaller reciprocal lattice vectors become relevant for the spin-polarized part of the tunneling current with coefficients that are consequently exponentially larger than those of the unpolarized part.

5 Examples of simulations and experiments

5.1 Seeing the Fermi surface in real space via the induced charge oscillations

Friedel oscillations define an important concept in quantum mechanics. They are created after perturbing an electron gas with an impurity. Charge and magnetic oscillations are then obtained in the surrounding electron gas. The shape and intensity of these oscillations contain important information on the impurity's electronic structure and on the band-structure of the host material where the impurity is embedded. As mentioned earlier, STM allowed to observe Friedel oscillations induced on surfaces characterized by two-dimensional electronic surface states. These surfaces, typically, Cu(111), Ag(111) and Au(111) surfaces provided the right playground for experimental and theoretical investigations on charge variations induced by impurities in a quasi-two-dimensional electron gas.

Isotropic Friedel oscillations

For Cu(111) the surface state shows a parabolic dispersion with a minimum at ~ 0.5 eV below the Fermi level. The corresponding band structure projected on the $\bar{\Gamma} - \bar{M}$ line of the

2D-Brillouin zone is shown in Fig. 8(a). The shaded regions indicate the regions in $E - \mathbf{k}_{\|}$ space, for which bulk eigenstates (Bloch waves) exist. Surface states can only exist in the white "gap"-regions. Two such states are indicated. Of special interest is the parabolic band with the minimum close to E_F, since this state is only partially occupied and gives rise to a two-dimensional metallic behavior, which is of great interest for the following.

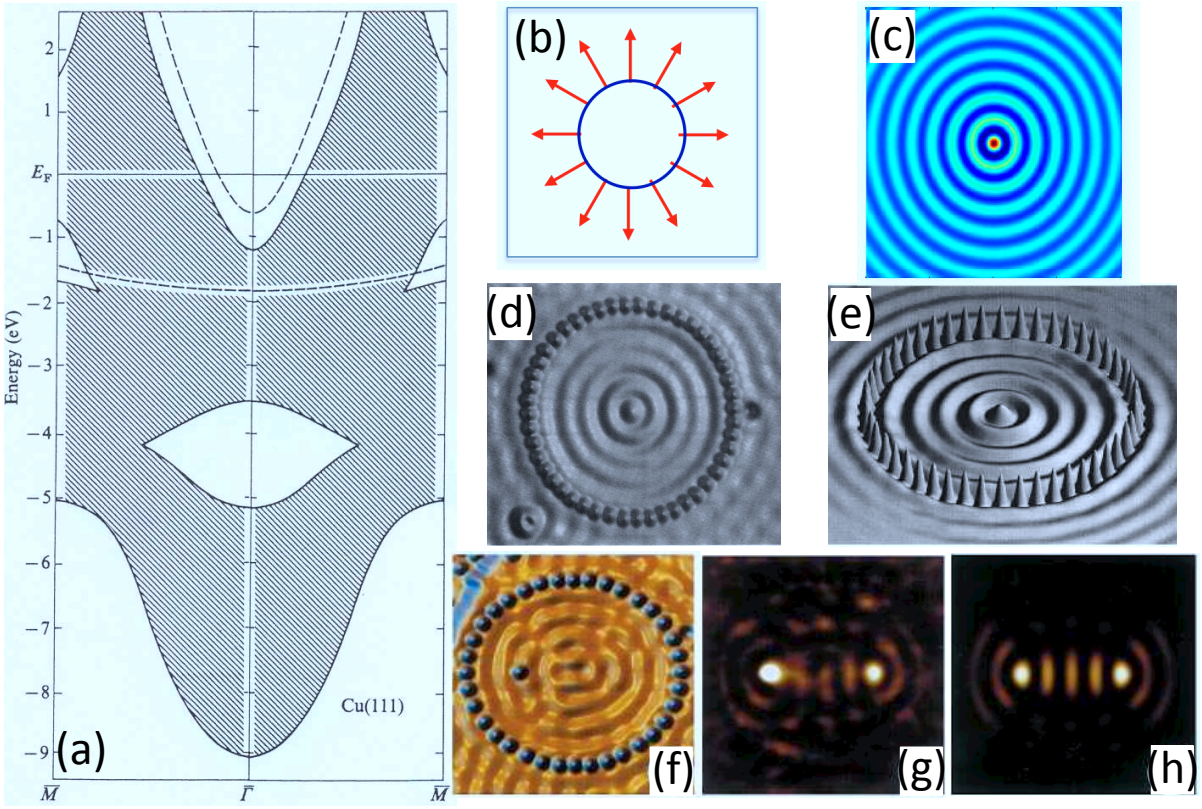


Fig. 8: *(a) Surface states (dashed curves) and bulk projected bands at a Cu(111) surface according to a six-layer surface band structure calculation [64]. In (b) and (c) illustration of the relation between the isotropic shape of the Fermi surface (black contours), group velocities (red arrows) in (b) and the corresponding induced isotropic charge oscillations in (c). In (d) and (e) is shown a comparison between the tunneling spectra obtained for a corral of Fe adatoms on Cu(111) surface as measured by Crommie et al. [5] (d) to the calculations made with the KKR method [65](e). Visualization of the quantum mirage with a mirage effect is shown in (f), (g) and (h). (f) is a topography showing an ellipse with $e = \frac{1}{2}$ and a Co atom at the left focus. In (g) the associated $\frac{dI}{dV}$ difference map shows the Kondo effect projected to the empty right focus, resulting in a Co atom mirage. This experimental measurement compares well with the calculated eigenmodes at E_F (magnitude of the wave function is plotted) as shown in (h). [7]*

For defects in the bulk, these Friedel oscillations of the charge perturbation vary for large distances r as $1/r^3$ times an oscillatory function and are in the jellium model proportional to:

$$\Delta n(r) \sim \frac{\cos(2k_F r + \delta)}{r^3} \tag{85}$$

However in the case of adatoms on surfaces, the charge response decays for long in-plane distances r slower than in the bulk and is determined by the surface states. In a free electron model, being well suited for the above surface state for Cu(111), the charge density is for large distances r proportional to

$$\Delta n(r) \sim \frac{\sin(2k_F r + \delta)}{r^2} \tag{86}$$

However, since also bulk states exist, which span most of the phase space (see Fig. 8(a)), the short range screening of the defect is dominated by these states, while only the long ranged behavior is determined by the surface state, which has a small wave vector k_F leading to long wave length oscillations.

In Fig. 8(b) and (c), are shown the Fermi surface of a two-dimensional electron gas with group velocities (vectors shown in red) and the corresponding induced charge around the impurity which is then isotropic and circular. The shape of the induced oscillations indicate that the related energy contour in reciprocal space is circular.

Many authors have observed such long ranged oscillations around adatoms, small clusters and steps on the Cu(111) surface in STM experiments. Most prominent among these is the work of the team of Eigler *et al.* [4, 5, 6, 7]. By atomic manipulations they were able to construct a corral of Fe atoms on the (111) Cu surface, and have shown that the surface states in the corral are more or less localized and form a discrete spectrum of resonant states. As an illustration of these we show in Fig. 8(d) and (e) a comparison of the experimental measurements to the result of calculations of Crampin *et al.* [65] obtained with the Korringa-Kohn-Rostoker Green function (KKR) method for a circular corral of 48 Fe atoms on the Cu(111) surface. Shown are the local density of states at the Fermi energy at 5 Å above the surface. Within the corral one sees a quantum well state with five maxima, corresponding to a localized state being more or less completely confined to the corral. Outside one sees oscillations arising from scattered surface state electrons at the corral, which decay with distance.

Let us shortly discuss the reason for the strong scattering of the surface state electrons at the Fe atoms. Basically in the vacuum region the full potential of Fe acts as a scattering center for the surface wave, being much stronger than the scattering at an Fe impurity in the bulk, where only the change of the Fe potential with respect to the host potential is effective. Moreover the wave vector k_F is relatively small, such that the wave length is considerably larger than the spacing between the Fe atoms. Therefore the surface wave does not "see" the corrugation of the Fe ring and is strongly reflected as in a cylindrical well. In fact the sequence of resonances can be well described by such a quantum well model, as has been shown recently [66]. The most fascinating corral experiments are the observation of atomic mirages in an elliptical quantum well [7]. An ellipse has the well known property that all classical waves emanating from one of the two focus points in every direction are reflected from the ellipse wall and focused in the second point, where these waves add up coherently since each such partial wave has the same path length and therefore the same phase shift. This is illustrated in Figs. 8(f–h) taken from Ref. [7]. Fig. 8(f) shows the STM topography for an ellipse with a given eccentricity, including one Co atom at the left focus point. Fig. 8(g) shows the $\frac{dI}{dV}$ difference maps, i.e. the change of the STM intensity map with respect to a small bias voltage V, which corresponds in the calculations to the local density of states in the vacuum region at the height of the STM tip. We see clearly two intensity spots, the real Co atom at the left focus and its image at the right focus. Thus in the empty focus we see the same accumulation of charge in the surface state as around the Co atom; therefore the image is called a quantum mirage. In fact the Co atom is a Kondo impurity

and a strong and sharp Kondo peak appears only in a very small energy region of about 10 meV around the Fermi level. Moreover the large mirage only appears, if one of the quantum well states falls into this energy region. Fig. 8(h) shows the calculated localized eigenstate observed in the experiment. The calculated local density of states compares very well with the $\frac{dI}{dV}$ curve shown in Fig. 8(h). Thus several conditions have to be satisfied for the Co mirage to appear:(i) the Co-atom has to sit in a focus point; if it sits at another position away from the focus point, no image appears, (ii) the bias voltage has to be such, that it coincides with an eigenstate of the ellipsoidal corral having maxima at the focus points, (iii) finally the image is particularly intense, if the eigenvalue coincides with the Kondo resonance. This concept has been recently extended theoretically by Stepanyuk and co-workers for the induced magnetization confined in magnetic corrals [67].

Focusing effect

If one manages to embed a Co-impurity few layers underneath the surface and try to visualize the induced Friedel oscillations on the surface, strange patterns are observed. Recently, we have shown by ab-initio calculations combined with STM observations that anisotropic localized oscillations can be observed on top of Cu(111) and Cu(001) surfaces due to the presence of buried Co impurities [68, 69]. These anisotropic ripples show that the usual isotropic free-electron model is not valid in such real situations. We demonstrated that these intriguing features are nothing else than a visualization in real space of parts of the bulk copper Fermi surface that are relatively flat. For the comparison between theory and STM, use is made of the Tersoff-Hamann model stating, as mentioned earlier, that scanning tunneling spectra can be related to the DOS in a certain energy interval in the vacuum.

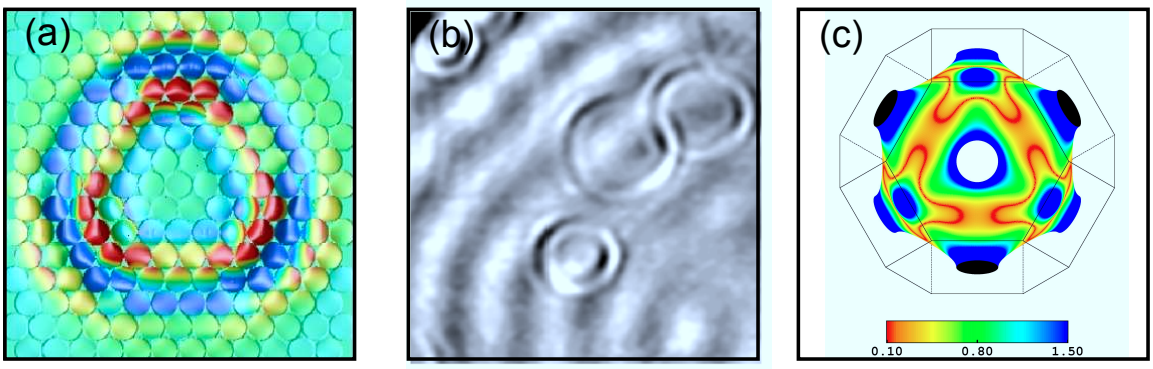


Fig. 9: *(a) Impurity induced charge density around E_F for an area of $\approx 30 \times 30\,\AA^2$ calculated at a height $\approx 6.1\,\AA$ above the Cu(111) surface with a Co impurity sitting in the 6^{th} layer below the surface. (b) Experimental STM topographies for an area of $90 \times 90\ \AA^2$ $(-80mV, 1nA)$ of four Co atoms below the Cu(111) surface. (c) Fermi surface of copper represented along the (111) direction. The inverse mass tensor corresponding to the denominator of Eq. 88 is represented by the color in units of the inverse electron mass. Small values represented in red lead to high intensities of the charge variation.*

Fig. 9(a) shows an example of the results of our simulations: the case of a Co impurity sitting

at the 6^{th} layer below the Cu(111) surface (-12.5Å below the surface). The charge induced in the vacuum has been computed up to an area of 30×30 Å^2 above the impurity. One notices the triangular shape of the induced charge with high values at the corners of the triangle. In addition, a one and a half period ripple can be observed to oscillate from the red positive value to the blue negative values and finally to the almost zero green values. The same period was also noticed in the STM experiment (Fig. 9(b)). To understand such a phenomenon, we start by giving the form of the induced variation of the DOS in the vacuum at position 0 by some buried Co impurity sitting at a position defined by $\vec{R}$ and inducing a change in the potential ΔV:

$$\Delta n_{\text{vacuum}}(\epsilon) \sim -\frac{1}{\pi}\Im\left[\int\int d^3r d^3r' G_0(0,\vec{r};\epsilon)\Delta V(\vec{r})G_0(\vec{r},0;\epsilon)\right] \tag{87}$$

where the Green function, G_0, obtained from the KKR method, describes the pure substrate. At very large distances R between the impurity and the vacuum site one can apply the stationary phase approximation and end up with a result similar to that of the well-known theory of interlayer exchange coupling:

$$R^2\Delta \text{n}_{\text{vacuum}} \propto \frac{1}{\left|\frac{d^2E}{dk_x^2}\cdot\frac{d^2E}{dk_y^2}\right|}. \tag{88}$$

The denominator of this equation is a measure of the curvature of the constant energy surface, i.e., the shape of the constant energy surface affects the propagation of the electrons. Additionally, one can show that the electronic waves are directed by the group velocities. Since, states at the Fermi energy (E_F) are probed experimentally, the constant energy surface corresponds to the Fermi surface: a small value of the curvature means that the Fermi surface has a flat region leading to large values of the DOS and to strong focusing of intensity in this space region determined by the group velocity. In Fig. 9(c), we show the Fermi surface, computed with ab-initio, of Cu oriented with the (111)-neck direction normal to the drawing plane. The Fermi surface is colored following the strength of the denominator of the right hand side of Eq. 88. The shape of the low values of this denominator, corresponding to the flat regions of the Fermi surface is found to be rather triangular along the (111) direction in accordance with our simulations of the induced Friedel oscillations. One can understand that the flat region seen within the triangle in Figs. 9(a) and (b) is induced by the neck of the Cu Fermi surface along the (111) direction that does not allow electrons to propagate.

Once the shape of the propagator G is known, either by STM-investigation or calculations based on density functional theory many additional effects can be predicted. The strong directionality of electron propagation even for a simple metal such as Cu, has consequences in many fields. For example, the spatially anisotropic characteristics should also be equally present in the RKKY interactions between magnetic impurities. Indeed we have shown that the interaction between Co adatoms on Pt(111) surface or Fe adatoms on Cu(111) surface is very anisotropic [20, 21], which is obviously induced by the same physics discussed in this subsection, i.e. the anisotropic shape of the Fermi surface.

5.2 Magnetism on surfaces with SP-STM

An example calculated by Heinze and collaborators [27] is shown in Fig. 10(a) and (b). A Cr overlayer on Ag(111) shows a row-wise antiferromagnetic structure (actually the ground state calculated by the authors is non-collinear, but for now we focus on the antiferromagnetic state).

The row-wise antiferromagnetism is indicated by the arrows drawn at the atomic positions, showing the magnetization direction of each atom. A scan of the surface with a non-magnetic STM tip will show the chemical unit cell (Fig. 10(a)): bright spots correspond to the Cr atoms, from which a tunneling current flows to the tip when the apex atom is above them. But a scan with a spin polarized tip reveals the magnetic structure (Fig. 10(b)). Instead of bright spots appearing around each atom, now bright stripes emerge at the rows with a magnetization direction parallel to the one of the tip, while dark stripes appear at the rows with opposite magnetization.
In these calculations, the assumption of a fully spin-polarized tip was made: i.e. $n_{T\downarrow}(E_F) = 0$. In Figs. 10(a) and (b), we see that the bright stripes (high current) appear when the substrate magnetization is parallel to the one of the tip. It should be noted that this is by no means guaranteed for all cases; it can well be that the antiparallel orientation favors the tunneling current. Eventually a case might be encountered where the parallel and antiparallel configurations give an almost equal signal at E_F. In this case spectroscopy is a very valuable tool, since one can detect the signal at other energies, choosing a voltage for which the two spin directions give considerably different results.
Another example of comparison between experiment and theory is shown in Figs. 10(c) and (d) for the case of Mn overlayer on W(110) surface. This work, realized by Heinze *et al.* [9], was the first observation of antiferromagnetism in a single magnetic layer. They found, as in the precedent example, that nonspin-polarized tunneling electrons image the chemical surface unit cell without any magnetic contribution, whereas spin-polarized electrons probe the change in translational symmetry due to the magnetic superstructure, which gives rise to a different image corresponding to the respective magnetic structure. The magnetic ground state that is antiferromagnetic in a checkerboard arrangement of Mn atoms with magnetic moments of opposite direction and an easy axis of the magnetization oriented in the film plane leads in the theoretical STM image to a stripe pattern similar to the one obtained experimentally. However, when the STM tip is considered non-magnetic, all Mn atoms become equivalent in the chemical unit cell and the STM-pattern becomes diamond-like both experimentally and theoretically. A detailed discussion on magnetism and in particular on magnetism at surfaces is given in Chapter **C 4** by S. Blügel.

5.3 Magnetic domain walls with TAMR-STM

The first experimental verification of the dependence of the DOS on the magnetization orientation was provided experimentally by Bode *et al.* [52] using a non spin-polarized STM on Fe double layers deposited on W(110) substrate. This effect discussed in Section 4.4, was already predicted theoretically [56]. The substrate chosen by Bode *et al.* is well known for having a nanometer-scale domain structure. In other words, the magnetization of the sample rotates at the nanometer-scale as shown in Fig. 11(a) and the idea is to probe the rotation of the magnetization at different positions using the spectroscopic mode, i.e. to measure the dI/dV spectra.

Interestingly, the domain walls are visible with a non-magnetic W tip (see Fig. 11(b)) along the different stripes propagating lateraly: At the position of the domain wall the differential conductivity dI/dV is reduced with respect to the domain. As revealed by the local tunneling spectra (Fig. 11(c)), this contrast is caused by a tiny difference which is energetically located just above E_F (see inset): while the dI/dV spectrum measured with the tip positioned above the domain exhibits a weak peak at a bias of 0.07 eV, this peak is almost absent in the domain wall

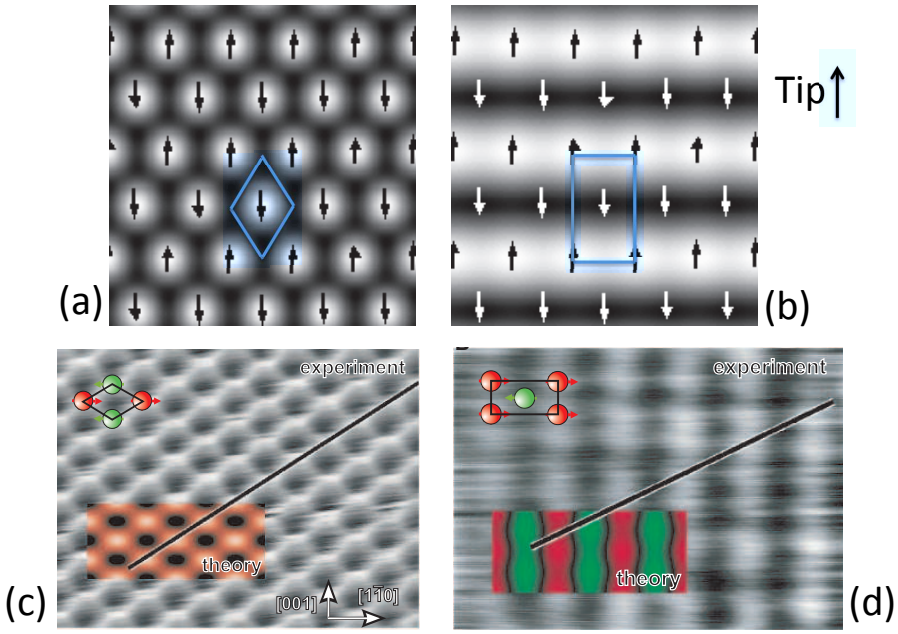


Fig. 10: *Calculated STM picture of an antiferromagnetic Cr monolayer on Ag(111) fcc surface after Ref. [27] and Mn monolayer on W(110) surface after Ref. [9]. In (a) shows the result obtained assuming a non spin-polarized tip, revealing the chemical unit-cell (drawn parallelogram); Cr atoms appear as white filled circles. In (b) a spin-polarized tip with magnetization direction in-plane as indicated. In this case the magnetic unit cell emerges (drawn as a rectangle), giving alternating black and white stripes. The drawn arrows indicate the magnetization direction of each Cr atom. A comparison of SP-STM measurements and first-principles calculations for the case of Mn layer on W(110) is shown in (c) considering a non-magnetic tip and in (d) when the tip is magnetic. Once more the unit-cell of the calculated magnetic ground state is shown in the insets of (c) and (d).*

spectra. This is further illustrated in the lower part of Fig. 11(c) by the plot of the TAMR ratio calculated as $[(dI/dV)_{\rm D} - (dI/dV)_{\rm W}]/[(dI/dV)_{\rm D} + (dI/dV)_{\rm W}]$, where the dI/dV are taken within the domain (D) or at the domaine wall (W).[3] For comparison, first-principles calculations of the LDOS and TAMR ratio on the same system including SOC were performed considering different rotations (out-of-plane and in-plane) of the magnetic moments are shown in Fig. 11(d). The agreement is good and it is found theoretically that the two minority-spin d_{z^2} states at the Fe surface lead to pronounced peaks at -0.18 and +0.85 eV and can be identified with the experimental peaks at -0.08 eV and +0.7 eV. A closer look (cf. inset in Fig. 11(d)) reveals a significant enhancement of the LDOS for the out-of-plane magnetized film within an energy interval of about 100 meV above E_F. The theoretical TAMR signal is qualitatively in line

[3] If the ratio is calculated on the original data (as measured), a pronounced oscillation can be found just below E_F. This oscillation is not caused by any additional or missing spectroscopic features in the domain wall of the dI/dV spectrum with respect to the spectrum measured at domains but by an overall energetic shift $\sim$ 11 meV. The physical origin is different work functions in domains and domain walls. After correction, the oscillation below E_F has almost perfectly disappeared.

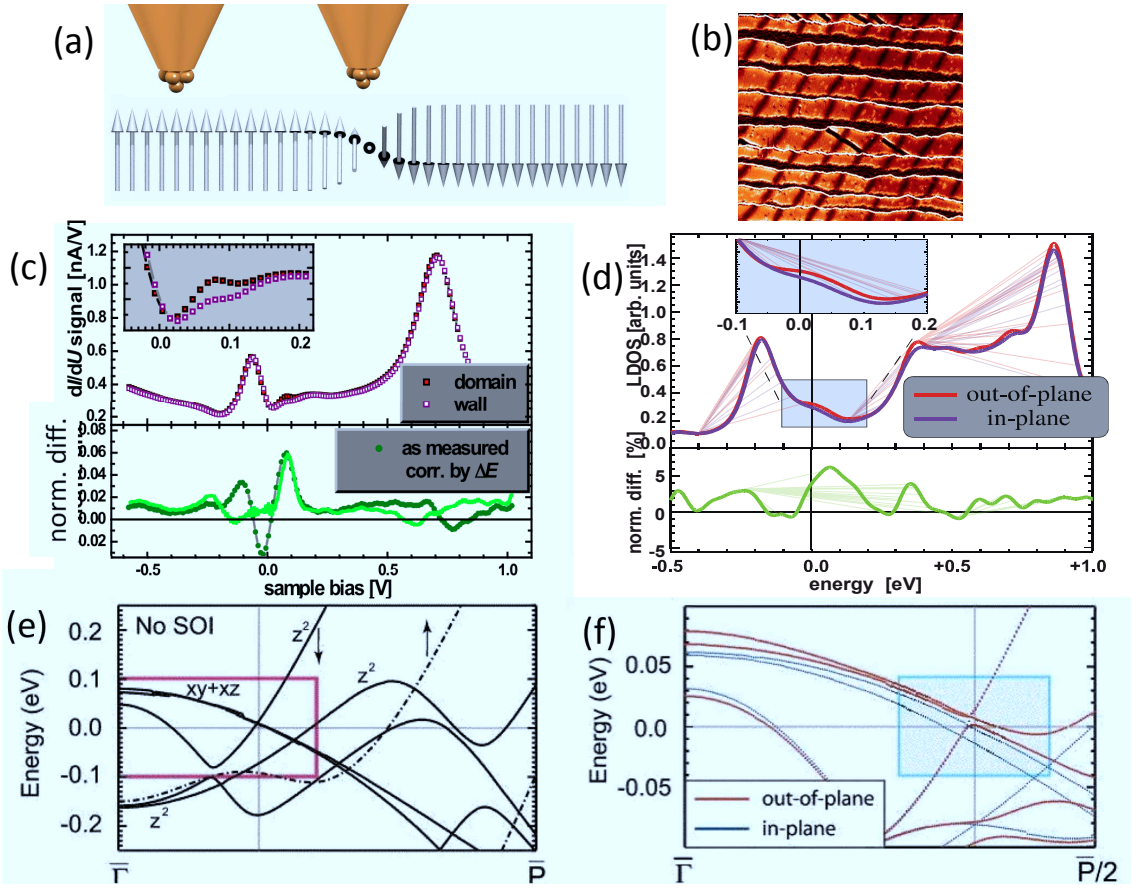


Fig. 11: *(a) Measurement of a domain wall with an STM-probe at different location: within the domain and at the domain-wall. (b) shows an image of the dI/dV signal measured on two-monolayers Fe deposited on W(110) substrate. Stripes of $\sim$ 50 nm widths can be observed with bright (within the domain $\sim$ in-plane orientation of the magnetization) and dark (at the domain-wall $\sim$ out-of-plane orientation of the magnetization) regions. In (c) and (d) are shown the bias-dependence of the dI/dV signal and the theoretical LDOS at different locations: domain (in-plane magnetization) and domain-wall (out-of-plane) regions. A TAMR ratio is computed and shown in lower parts of (c) and (d). The band-structure calculated without (e) and with SOC (f) demonstrates the impact of SOC and of the orientation of the magnetization on the electronic structure of the sample. Figures taken and adapted from Ref.[52]*

with the experimental signal: The peak at -0.3 eV could be identified with the experimental peak at -0.24 eV. The origin of the dependence of the LDOS on SOC can be traced back to the change of the band-structure once the magnetization is rotated as discussed in Section 4.4. The band-structure without and with SOC are plotted in Figs. 11(e) and (f). States at the $\overline{\Gamma}$ point decay the lowest into vacuum (as discussed in Section 4.5). Around -0.18 eV and +0.05 eV, the d_{z^2}-band crosses the $\overline{\Gamma}$-point. Because of their appropriate extension in the z-direction, the d_{z^2} orbitals decay slower into the vacuum compared to the d_{xz} and d_{yz} orbitals. Interestingly, the crossing of the minority d bands that occured without SOC is avoided once SOC is included

and if the magnetization is rotated out-of-plane (red line in Fig. 11(f)). This has a strong impact by shifting, for example, the crossing of the bands at the $\overline{\Gamma}$-point. Overall such a small effect is enough to induce a change in the LDOS that is observable with STM. We note that the use of TAMR within STM was recently extended to probe adatoms deposited on surfaces with magnetic domain walls [54, 55].

Acknowledgment

I would like to acknowledge the HGF-YIG Programme VH-NG-717 (Functional Nanoscale Structure and Probe Simulation Laboratory–Funsilab) for financial support.

References

[1] G. Binnig, H. Rohrer, Ch. Gerber and E. Weibel, Appl. Phys. Lett. **40**, 178 (1982).

[2] G. Binnig, H. Rohrer, Ch. Gerber and E. Weibel, Phys. Rev. Lett. **49**, 57 (1982).

[3] G. Binnig, H. Rohrer, Ch. Gerber and E. Weibel, Phys. Rev. Lett. **50**, 120 (1983).

[4] D. M. Eigler and E. K. Schweizer, Nature **344**, 524 (1990).

[5] M. F. Crommie, C. P. Lutz, and D. M. Eigler, Science **262**, 218 (1993).

[6] E. J. Heller, M. F. Crommie, C. P. Lutz and D. M. Eigler, Nature **369**, 466 (1994).

[7] H. C. Manoharan, C. P. Lutz and D. M. Eigler, Nature **403**, 512 (2000).

[8] R. Wiesendanger, H.-J. G”untherodt, G. G”untherodt, R. J. Gambino and R. Ruf, Phys. Rev. Lett. **65**, 247 (1990).

[9] S. Heinze, M. Bode, A. Kubetzka, O. Pietzsch, X. Nie, S. Blügel and R. Wiesendanger, Science **288**, 1805 (2000).

[10] S. Blügel, M. Weinert and P. H. Dederichs, Phys. Rev. Lett. **60**, 1077 (1988).

[11] M. Bode, M. Heide, K. von Bergmann, P. Ferriani, S. Heinze, G. Bihlmayer, A. Kubetzka, O. Pietzsch, S. Blügel and R. Wiesendanger, Nature **447**, 190 (2007).

[12] S. Loth, S. Baumann, C. P. Lutz, D. M. Eigler and A. J. Heinrich, Science **335**, 196 (2012).

[13] S. Holzberger, T. Schuh, S. Blügel, S. Lounis and W. Wulfhekel, Phys. Rev. Lett. **110**, 157206 (2013).

[14] B. C. Stipe, M. A. Razaei and W. Ho, Science, **279**, 5358 (1998).

[15] A. J. Heinrich, J. A. Gupta, C. P. Lutz and D. M. Eigler, Science **306**, 5695 (2004).

[16] C. F. Hirjibehedin, C. P. Lutz, A. J. Heinrich, Science, **312**, 5776 (2006).

[17] T. Balashov, T. Schuh, A. F. Takács, A. Ernst, S. Osctanin, J. Henk, I. Merting, P. Bruno, T. Miyamachi, S. Suga and W. Wulfhekel, Phys. Rev. Lett. **102**, 257203 (2009).

[18] A. A. Khajetoorians, S. Lounis, B. Chilian, A. T. Costa, L. Zhou, D. L. Mills, J. Wiebe and R. Wiesendanger, Phys. Rev. Lett. **106**, 037205 (2011).

[19] S. W. Wu, N. Ogawa and W. Ho, Science **312**, 5778 (2003).

[20] L. Zhou, J. Wiebe, S. Lounis, E. Vedmendenko, F. Meier, S. Blügel, P. H. Dederichs, R. Wiesendanger, Nature Physics **6**, 187 (2010).

[21] A. A. Khajetoorians, J. Wiebe, B. Chilian, S. Lounis, S. Blügel, R. Wiesendanger, Nature Physics **8**, 497 (2012).

[22] C.J. Chen, *Introduction to Scanning Tunneling Microscopy* (Oxford University Press, Oxford, 1993).

[23] J. Tersoff and D.R. Hamann, Phys. Rev. Lett. **50**, 1998 (1983).

[24] J. Tersoff and D.R. Hamann, Phys. Rev. B **31**, 805 (1985).

[25] C.J. Chen, Phys. Rev. B **42**, 8841 (1990).

[26] C. J. Chen, Phys. Rev. Lett. **65**, 448 (1990).

[27] S. Heinze, P. Kurz, D. Wortmann, G. Bihlmayer, G. Bihlmayer and S. Blügel, Appl. Phys. A **75**, 25 (2002).

[28] N. Lorente and M. Persson, Phys. Rev. Lett. **85**, 2997 (2000).

[29] N. Lorente and J. P. Gauyacq, Phys. Rev. Lett. **103**, 176601 (2001).

[30] M. Persson, Phys. Rev. Lett. **103**, 050801 (2009).

[31] J. Fernandez-Rossier, Phys. Rev. Lett. **102**, 256802 (2009).

[32] J. Fransson, Nano Lett. **9**, 2414 (2009).

[33] A. Hurley, N. Baadji, and S. Sanvito, Phys. Rev. B **86**, 125411 (2012).

[34] S. Lounis, A. T. Costa, R. B. Muniz and D. L. Mills, Phys. Rev. Lett. **105**, 187205 (2010).

[35] S. Lounis, A. T. Costa, R. B. Muniz and D. L. Mills, Phys. Rev. B **83**, 035109 (2011).

[36] H.-J. Günterodt and R. Wiesendanger (Ed.), *Scanning Tunneling Microscopy I–III*, (Springer, Heidelberg, 1994–1996).

[37] R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications* (Cambridge University Press, Cambridge, 1994).

[38] S. Heinze, *First-Principles Theory of Scanning Tunneling Microscopy Applied to Transition-Metal Surfaces*, PhD thesis, University of Hamburg (2000).

[39] W. A. Hofer, A. S. Foster and A. L. Schluger, Rev. Mod. Phys. **75**, (2003).

[40] M. Bode, Rep. Prog. Phys. **66**, 523 (2003).

[41] R. Wiesendanger, Rev. Mod. Phys. **81**, 1495 (2009).

[42] A. van Housel and H. J. W. Zandvliet, Rev. Mod. Phys. **82** 1593 (2010).

[43] J. P. Gauyacq, N. Lorente and F. D. Novaes, Prog. Surf. Sci. **87**, 63 (2012).

[44] S. Blügel, *Theorie der Rastertunnelimikroskopie*, in the lecture notes of the 30^{th} IFF Spring School, Forschungszentrum Jülich (1999).

[45] K. K. Saha, J. Henk, A. Ernst, and P. Bruno, Phys. Rev. B **77**, 085427 (2008).

[46] C. F. Quate, *Vacuum Tunneling: A New Technique for Microscopy*, Physics Today, 26 (1986).

[47] J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).

[48] M. Jullière, Physics Letters A, **54** 225 (1975).

[49] J. Slonczewski, Phys. Rev. B, **39**, 6995 (1989).

[50] J. S. Moodera, L. R. Kinder, T. M. Wong and R. Meservey, Phys. Rev. Lett. **74**, 3273 (1995).

[51] T. Yaoi, S. Ishio and T. Miyazaki, J. Mag. Mag. Mat. **126**, 430 (1993).

[52] M. Bode, S. Heinze, A. Kubetzka, O. Pietzsch, X. Nie, G. Bihlmayer, S. Blügel and R. Wiesendanger, Phys. Rev. Lett. **89**, 237205 (2002).

[53] C. Gould, C. Rüster, T. Jungwirth, E. Girgis, G. M. Schott, R. Giraud, K. Brunner, G. Schmidt and L. W. Molenkamp, Phys. Rev. Lett. **93**, 117203 (2004).

[54] N. Néel, S. Schröder, N. Ruppelt, P. Ferriani, J. Kröger, R. Berndt and S. Heinze, Phys. Rev. Lett. **110**, 037202 (2013).

[55] D. Serrate, P. Ferriani, Y. Yoshida, S-W. Hla, M. Menzel, K. von Bermann, S. Heinze, A. Kubetzka, R. Wiesendanger, Nature Nanotechnology, **5**, 350 (2010).

[56] A. Lessard, T. H. Moos, and W. Hübner, Phys. Rev. B **56**, 2594 (1997).

[57] T. Balashov, A. F. Takács, M. Däne, A. Ernst, P. Bruno, W. Wulfhekel, Phys. Rev. B **78**, 1774404 (2008).

[58] S. Loth, M. Etzkorn, C. P. Lutz, D. M. Eigler, A. J. Heinrich, Science **329**, 5999 (2010).

[59] A. A. Khajetoorians, B. Baxevanis, C. Hübner, T. Schlenk, S. Krause, T. O. Wehling, S. Lounis, A. Lichtenstein, D. Pfannkuche, J. Wiebe and R. Wiesendanger, Science **339**, 6115 (2013).

[60] T. Miyamachi, T. Schuh, T. Markl, C. Bresch, T. Balashov, A. Stohr, C. Karlewski, S. Andre, M. Marthaler, M. Hoffmann, M. Geilhufe, S. Ostanin, W. Hergert, I. Mertig, G. Schon, A. Ernst and W. Wulfhekel, Nature **503**, 242 (2013).

[61] D. Wortmann, S. Heinze, Ph. Kurz, G. Bihlmayer and S. Blügel, Phys. Rev. Lett. **86**, 4132 (2001).

[62] S. Schröder, *First-Principles Study of Non-Collinear Magnetism and Spin-Orbit Driven Physics in Nanostructures at Surfaces*, PhD thesis, Christian-Albrechts University in Kiel (2013).

[63] E. Abate and M. Asdente, Phys. Rev. B **140**, A1303 (1965).

[64] A. Euceda, D. M. Bylander and L. Kleinman, Phys. Rev. B **28**, 528 (1983).

[65] S. Crampin and O. R. Bryant, Phys. Rev. B **54**, 17367 (1996).

[66] B. Lazarovits, B. Újfalussy, L. Szunyogh, B. L. Györffy and P. Weinberger, J. Phys.: Condens. Matter **17**, S1037 (2005).

[67] V. S. Stepanyuk, L. Niebergall, W. Hergert and P. Bruno, Phys. Rev. Lett. **94**, 187201.

[68] A. Weismann, M. Wenderoth, S. Lounis, P. Zahn, N. Quaas, R. G. Ulbrich, P. H. Dederichs and S. Blügel, Science **323**, 1190 (2009).

[69] S. Lounis, P. Zahn, A. Weismann, M. Wenderoth, R. G. Ulbrich, I. Mertig, P. H. Dederichs and S. Blügel, Phys. Rev. B **83**, 035427 (2011).

C 7 Transport through correlated nanostructures[1]

Maarten R. Wegewijs
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

The aim of this chapter is to illustrate the application of the real-time approach to open quantum systems presented in lecture B5 by H. Schoeller to transport problems for a range of correlated nanoscale systems. Here we focus on a particular, yet quite broad class of correlated nanoscale systems referred to as "quantum dots". A quantum dot is a zero-dimensional system, i.e., the electronic wave functions are confined in all spatial directions, and the resulting energy gaps can be resolved experimentally at low temperatures (few K and below). The excitations responsible for interesting transport features are due to various basic effects, e.g., magnetic field, vibrational excitations, or spin-orbit interaction, and more complicated interplay, e.g., between orbitals and vibrations or spin and vibrations. However, the dominating energy scale is that of the Coulomb interaction and this makes the theoretical description challenging, but also more interesting than just a theory of transport spectroscopy: due to the interplay of interaction and quantum fluctuation effects the quantum dot properties can be renormalized, an effect which already appears in the *leading* order of perturbation theory. As a result, the transport shows pronounced effects that do not directly relate to the "bare" spectrum of the system which is studied using electron tunneling current. How to deal with this in calculations that are perturbative in the tunnel coupling is already a delicate matter and it is these issues that are at the focus here, thereby complementing lecture B5. To achieve this, the chapter is organized according to two theoretical principles, the order of perturbation theory and the importance of density-operator coherences in the energy basis (see below). We illustrate the importance of the different possible cases with applications to experiments on various correlated nanostructures, in particular, organic molecules [Sec. 3.1], carbon nanotubes, and single-molecule magnets [both in Sec. 4.2]. Before that, we start with a brief summary of results from the real-time approach and establish some notation [Sec. 2]. Extensive treatment of the material presented here can be found in [1, 2] and the experimental works cited below. A modern development, the introduction of second quantization techniques in Liouville space, the so-called "superfermion" approach, is described in detail in [3, 4]. Not covered here, is the application of an earlier formulation [5] of the RG approach presented in lecture B5 to extend the real-time perturbation theory developed here [3]. The chapter does not aim to review the broad area of transport through correlated nanoscale systems for which reviews are available, some recent examples being [6, 7, 8]. Also, no attempt is made to do justice to the long history of some topics (e.g., the SET theory dating back to 1975 [9]), which is covered in, e.g., [10].

2 Density operator and real-time diagrammatic approach

The real-time approach to the dynamics of open quantum systems was introduced in lecture B5 in a much broader context. Here we focus on time-independent systems driven into a stationary state by bilinear electron tunnel coupling [i.e., single vertex operators, cf. Eq. (3)] in the wide-band limit (see below). For orientation we point out that the real-time approach, formulated initially using Keldysh contours [11] and later using Liouville space superoperators [5, 1] is a (reduced) *density operator approach* [12]. It is exactly equivalent [13, 2] to the much older Nakajima-Zwanzig projection approach[2] to quantum dynamics which is in use today in various fields of physics and chemistry. However, to this the real-time approach adds *quantum-field theoretical techniques* which are advantageous in concrete calculations. By introducing Wick's

[2] More precisely, with the specific projector $P\bullet = \rho^{\rm res} \otimes \mathop{\rm Tr}_{\rm res} \bullet$ in the notation of the next section.

theorem for open quantum systems [5, 3, 4] from the start, and by using diagrams to set up the perturbation theory and formulate approximations, one can maintain detailed insight and overview without having to write down cumbersome expressions. This is not only important for advanced approximation schemes, such as the diagrammatic real-time renormalization group approach of lecture B5, but also in perturbative calculations beyond the leading order [2] that we consider here. In the following we set $\hbar = k_{\rm B} = |e| = 1$.

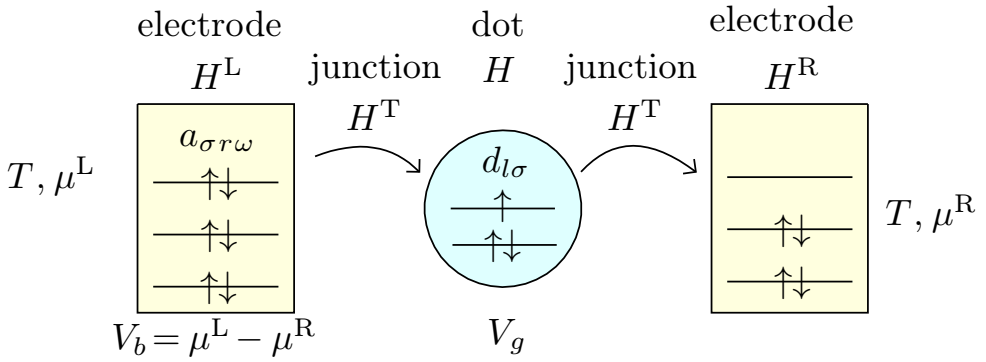


Fig. 1: *Quantum dot with orbitals labeled l and spin σ coupled to electrodes with a continuum of levels. The single- and many-particle levels of the dot are uniformly and linearly tuneable with the gate voltage V_g. In general they also depend on the bias $V_b = \mu^L - \mu^R$ due to capacitive effects [14, 15] unless the capacitive coupling is symmetric, which we assume here for simplicity. See, however, the closing remarks of this chapter.*

2.1 General transport model

We consider a rather general model Hamiltonian $H^{\rm tot}$ for a quantum dot tunnel coupled to electrodes labeled by $r = {\rm L}, {\rm R}$, as sketched in Fig. 1:

$$H^{\rm tot} = H + \textstyle\sum_r H^r + H^{\rm T}, \quad \begin{cases} H = \displaystyle\sum_a E_a |a\rangle\langle a| & \text{dot,} \quad (1) \\ H^r = \displaystyle\sum_{\sigma'} \int d\omega\, (\omega + \mu^r)\, a^\dagger_{\sigma' r\omega} a_{\sigma' r\omega} & \text{electrodes,} \quad (2) \\ H^{\rm T} = \displaystyle\sum_r \sum_{\sigma' \sigma l} \int d\omega\, t_{\sigma l \sigma' r\omega}\, d^\dagger_{\sigma l} a_{\sigma' r\omega} + \text{h.c} & \text{junctions.} \quad (3) \end{cases}$$

We are interested in strongly correlated quantum dots (H) with arbitrary many-body eigenstates $|a\rangle$ and many-body energies E_a. Electron field operators on the dot are denoted $d_{\sigma l}$ with spin σ and an some orbital quantum number l. Besides electronic degrees of freedom, the quantum dot states $|a\rangle$ may contain other quantum numbers, e.g., describing quantized vibrations [see Eq. (65)]. The noninteracting electrodes (H^r) with fields $a_{\sigma r\omega}$ have a continuous spectrum with energy ω measured relative to the electrochemical potential μ^r of the respective electrode with temperature T. The tunnel junction ($H^{\rm T}$) can be quite general, in particular, the spin dependence included in $t_{\sigma l\sigma' r}$ will be important in Sec. 3.2. Here, the spin-quantization axes are chosen independently in the dot and the electrodes (σ and σ' refer to different quantization axes). The main approximation is the wide-band limit, i.e., the assumption that $t_{\sigma l\sigma' r\omega}$, already including the density of states, is a constant for frequencies $|\omega| \leq D$ and zero otherwise, and

that the band-width $2D$ is by far the largest energy scale. The electron number and current operator are given by [B5, Eqs. (70), (78)], respectively,

$$N^r = \sum_{\sigma'} \int d\omega a^\dagger_{\sigma' r\omega} a_{\sigma' r\omega}, \qquad I^r = -\widehat{\frac{dN^r}{dt}} = -i[H^{\mathrm{T}}, N^r]. \tag{4}$$

We make the quantum statistical assumption that before the tunnel coupling H^{T} is switched on at time $t = 0$, the dot and reservoirs are not correlated [cf. Eq. (9) below]. Moreover, the density operator describing the reservoirs is the (tensor)product of the separate equilibrium states of each reservoir r with electrochemical potential μ^r and temperature T: with $\underset{r}{\mathsf{Tr}}$ denoting the trace over electrode r

$$\rho^{\mathrm{res}} = \prod_r \frac{1}{Z^r} e^{-(H^r - \mu^r N^r)/T}, \qquad Z^r = \underset{r}{\mathsf{Tr}} e^{-(H^r - \mu^r N^r)/T}. \tag{5}$$

2.2 Kinetic / quantum master equation

The time evolution of the density operator of the total system, dot plus electrodes, is described by the Liouvillian superoperator (= a linear map on operators), given by the commutator of the Hamiltonian with an operator argument denoted by $\bullet$:

$$L^{\mathrm{tot}} = L + \textstyle\sum_r L^r + L^{\mathrm{T}}, \quad \begin{cases} L\bullet = [H, \bullet]_- & \text{dot,} \quad (6) \\ L^r\bullet = [H^r, \bullet]_- & \text{electrodes,} \quad (7) \\ L^{\mathrm{T}}\bullet = [H^{\mathrm{T}}, \bullet]_- & \text{junctions.} \quad (8) \end{cases}$$

The solution of the Liouville von-Neumann equation

$$\frac{d}{dt}\rho^{\mathrm{tot}}(t) = -iL^{\mathrm{tot}}\rho^{\mathrm{tot}}(t), \qquad \rho^{\mathrm{tot}}(0) = \rho^{\mathrm{res}} \otimes \rho(0), \tag{9}$$

where $\rho(0)$ is an arbitrary initial state of the dot, reads as

$$\rho^{\mathrm{tot}}(t) = e^{-iL^{\mathrm{tot}}t}\rho^{\mathrm{res}} \otimes \rho(0) \neq \rho^{\mathrm{res}} \otimes \rho(t). \tag{10}$$

As indicated, this state cannot be factorized into the equilibrium reservoir part and some dot density operator $\rho(t)$ for times $t > 0$ due the tunnel coupling L^{T}. This holds in particular when going *beyond* the leading nonvanishing order $\sim (L^{\mathrm{T}})^2$, as we will do here, and failing to systematically account for this leads to problems, see [2]. Instead, we need to calculate the (reduced) density operator, defined as

$$\rho(t) := \underset{\mathrm{res}}{\mathsf{Tr}}\rho^{\mathrm{tot}}(t), \tag{11}$$

from which all quantities of interest can be found. As shown in lecture B5 this operator and the expectation for the current obey [B5, Eqs. (2), (23), (69), and (76)]

$$\frac{d}{dt}\rho(t) = -iL\rho(t) + \int_0^t dt' W(t,t')\rho(t'), \qquad W(t,t') := -i\Sigma(t,t'), \tag{12}$$

$$\langle I^r\rangle(t) = \mathsf{Tr}\int_0^t dt' W^{I^r}(t,t')\rho(t'), \qquad W^{I^r}(t,t') := -i\Sigma^{I^r}(t,t'), \tag{13}$$

where Tr denotes the trace over the quantum dot. Here we introduced time-evolution and current *rate* superoperators $W(t,t')$ and $W^{I^r}(t,t')$ [rather than self-energies $\Sigma(t,t')$ and $\Sigma^{I^r}(t,t')$] to facilitate the following discussion. Equation (12) is called kinetic equation or quantum master equation.

Stationary state. In this chapter we are interested only in (i) time-independent Hamiltonians / Liouvillians, (ii) the stationary (long-time) limit $t \to \infty$ of the transport current expectation value, and (iii) the wide-band limit mentioned above. We therefore perform the half-sided Fourier transform of the kinetic equation (12), $\rho(E) = \int_0^\infty dt e^{iEt}\rho(t)$, and solve it by a Green or resolvent superoperator [B5, Eq. (27)],

$$\rho(E) = \frac{i}{E - L - iW(E)}\rho(0). \tag{14}$$

With the subsidiary condition that the stationary state is trace-normalized, $\mathsf{Tr}\rho(\infty) = 1$, the central equations are then found to be [B5 Eqs. (80)-(81)],

$$0 = [-iL + W(i0)]\ \rho(\infty) \qquad \text{stationary state,} \tag{15}$$

$$\langle I^r\rangle(\infty) = \mathsf{Tr}\, W^{I^r}(i0)\,\rho(\infty) \qquad \text{stationary current.} \tag{16}$$

As expected, the long time limit of the density operator and the current, $\rho(\infty)$ and $\langle I^r\rangle(\infty)$, are determined by the respective zero-frequency rate superoperators, $W(i0)$, and $W^{I^r}(i0)$. Since we focus entirely on the stationary limit we will drop the arguments $(i0)$ and (∞) from hereon.

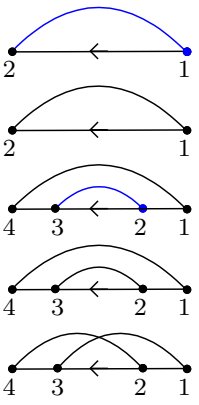
Perturbation theory. The results summarized so far apply quite generally to density operator approaches and can also be obtained within, e.g., the Nakajima-Zwanzig projection approach. The key advantages of the real-time approach appear when evaluating the rate superoperators required above. In this chapter we will review the results for the nonlinear transport when W and W^{I^r} are evaluated in the leading order in the tunnel coupling, $\Gamma \sim (L^{\mathrm{T}})^2 \sim (H^{\mathrm{T}})^2$, plus next-to-leading order:

$$W \approx W^{(1)} + W^{(2)}, \tag{17}$$

where $W^{(1)} \sim \Gamma^1$ and $W^{(2)} \sim \Gamma^2/T$. For $T \to \infty$ the exact result is contained in $W^{(1)}$ [3, 4] and the finite temperature corrections contained in both $W^{(1)}$ and $W^{(2)}$ are of interest. Large prefactors in terms of higher order in Γ may cause the perturbative approach (17) to break down and then the RG methods of lecture B5 are required [3]. Here we focus on regimes where this approach (17) does work due to sufficiently high T. The evaluation of W^{I^r} proceeds similarly [B5, Eqs. (67) and (77)] and will not be discussed here [1]. The diagrammatic rules [3] allow one to write down the rate superoperator directly as a sum of irreducibly contracted superoperators with intermediate "free" resolvents [B5, Eq. (63)]:[3]

[3] In the Nakajima-Zwanzig projection approach the complement to the projection operator P (see previous footnote) forces expressions to be irreducible, and via Wick's theorem one ends up with same result, see [13, 2, 16].

$$
\begin{aligned}
iW^1 \;=\;& G_2^+ \frac{\gamma^{\mathrm{R}}_{21}}{i0-L-x_1} G_1^- \\
&+G_2^+ \frac{\gamma^{K}_{21}(x_1)}{i0-L-x_1} G_1^+ \\
iW^2 \;=\;& G_4^+ \frac{1}{i0-L-x_1} G_3^+ \frac{\gamma^{\mathrm{R}}_{32}}{i0-L-x_1-x_2} G_2^- \frac{\gamma^{\mathrm{K}}_{41}(x_1)}{i0-L-x_1} G_1^+ \\
&+G_4^+ \frac{1}{i0-L-x_1} G_3^+ \frac{\gamma^{\mathrm{K}}_{32}(x_2)}{i0-L-x_1-x_2} G_2^+ \frac{\gamma^{\mathrm{K}}_{41}(x_1)}{i0-L-x_1} G_1^+ \\
&-G_4^+ \frac{}{i0-L-x_2} G_3^+ \frac{\gamma^{\mathrm{K}}_{42}(x_2)}{i0-L-x_1-x_2} G_2^+ \frac{\gamma^{\mathrm{K}}_{31}(x_1)}{i0-L-x_1} G_1^+
\end{aligned}
$$



(18)

Importantly, the reservoirs have been *explicitly* eliminated: the Liouvillian $L\bullet = [H,\bullet]_-$ clearly only acts on the dot subsystem, as do the remaining superoperators $G_1^\pm$:

$$
G_1^q \bullet = \tfrac{1}{\sqrt{2}} \begin{cases} d_1 \bullet + (-1)^N \bullet (-1)^N d_1 & q=+ \\ d_1 \bullet - (-1)^N \bullet (-1)^N d_1 & q=- \end{cases}, \qquad d_1 \bullet = \begin{cases} d^\dagger_{\sigma_1 l_1} & \eta_1 = + \\ d_{\sigma_1 l_1} & \eta_1 = - \end{cases}, \tag{19}
$$

Here $N = \sum_{\sigma l} d^\dagger_{\sigma l} d_{\sigma l}$ is the dot number operator and multiindices $i = \eta_i \sigma_i l_i r_i$ with $i = 1,2,3,4$ were used. The admittedly non-intuitive expressions (19) are in fact *quantum field superoperators* and have properties very similar to the fermionic quantum field *operators* d_1 from which they are constructed.[4]

This allows many second quantization techniques to be employed also for *open* quantum systems and reveals fundamental aspects of the structure of their dynamics [3, 4]. The only nontrivial thing still left from the reservoirs are the integrations over the reservoir frequencies $x_i = \eta_i(\omega_i + \mu_i)$ which also appear in the contraction functions

$$
\gamma^{\mathrm{R}}_{21} = \Gamma_{21} \cdot 1 \qquad \text{retarded contraction,} \tag{20a}
$$

$$
\gamma^{\mathrm{K}}_{21} = \Gamma_{21} \cdot \tanh\left(\frac{x_2 - \eta_2 \mu_2}{2T}\right) \qquad \text{Keldysh contraction,} \tag{20b}
$$

where $\mu_i := \mu^{r_i}$. The spectral density function,

$$
\Gamma_{21} := \Gamma_{\eta_2\sigma_2 l_2 r_2, \eta_1 \sigma_1 l_1 r_1} := \delta_{\eta_2,-\eta_1} \delta_{r_2,r_1} \sum_{\sigma_3} t_{\eta_2 \sigma_2 l_2, \sigma_3 r_2} (t_{\eta_1 \sigma_1 l_1, \sigma_3 r_1})^* \tag{21}
$$

with $i = \eta_i \sigma_i l_i r_i$. and $t_{+\sigma_2 l_2 \sigma_3 r_3} = (t_{-\sigma_2 l_2 \sigma_3 r_3})^* := t_{\sigma_2 l_2 \sigma_3 r_3}$ [cf. Eq. (3)], can be considered as the "tunnel rate" for an elementary process (contraction), but also contains interference information (it is complex). The contraction functions (20) arise due to the application of the Wick theorem [3] and in the factors 1 and $\tanh(x/2)$ one may recognize the reservoir correlation functions, in particular the retarded and Keldysh component discussed in lecture B2 by S. Jakobs [B2, Eqs. (39), (144) and (148)]. [5] The rate superoperator W thus has a physi-

[4] These correspond to the vertex operators of lecture B5 *summed / subtracted* with respected to their Keldysh indices p [cf. B5, Eq. (92)].

[5] These are the (anti)symmetric part of the Fermi function $f(x) = (e^x + 1)^{-1}$ [cf. B5, Eq. (91)], and their appearance has fundamental origins. The fact that the retarded contraction function (20a) depends on the frequency *only* through Γ_{21} has important consequences for the perturbation theory, and not only in the more involved real-time RG approach [cf. B5, Eq. (91)], in particular in the wide band limit where Γ_{21} is independent of the frequency $\omega_2 = \omega_1$. Also, the fact that the advanced correlation function does *not* appear in the Liouville-space field theory imposes a strong restriction on the perturbation series ("causal structure") which is not recognized in other formulations [11] of the real-time approach, or other density operator approaches [3, 4].

cally appealing form: (the irreducible part of) the time evolution is expressed as a sequence of tunnel processes ($G^{\pm}$), with intermediate free time evolution (resolvents), the processes being weighted by the correlation function of the reservoirs. The only actual calculation consists of (i) taking supermatrix elements (see below) of Eq. (18) in the quantum dot the energy basis $\{|a\rangle\}$ $[H|a\rangle = E_a|a\rangle]$ and (ii) performing the integrations over the reservoir frequencies x_i and (iii) summing the expression over all the abbreviated indices $i = \eta_i\sigma_i l_i$ with $i = 1, 2, 3, 4$. In the subsequent Sections we will occasionally comment on some of the above details, but only when they relate to the physics contained in an expression.
For now, the key point to remember is that W can be calculated systematically in powers of Γ and that we include both the first and the second order (i.e., up to the fourth order in L^{T}). A complete account of the evaluation of $W^{(1)} + W^{(2)}$ can be found in [1] (using a related but somewhat earlier formulation). A further detailed analysis, exploiting relations between the Liouville-space and Keldysh-diagram techniques, is given in [2] and reveals simplifications important for numerical implementation of the scheme (18), as well as the precise relation to the so-called "Golden Rule + T-matrix" approach, pin-pointing dangerous flaws in the way that the latter approach is usually applied.

3 First-order tunneling effects

Having outlined how the rate superoperator W can be calculated to second order, we now discuss the physical content of this approximate theory. We start with the first-order approximation (often called "Born approximation"),

$$W \approx W^{(1)} \sim \Gamma, \tag{22}$$

and consider the solution of the equation (15) for the stationary density operator:

$$\frac{d}{dt}\rho = 0 = \left(-iL + W^{(1)}\right)\rho. \tag{23}$$

To exploit the structure of the superoperator L, we decompose the density operator in the energy eigenbasis of the dot Hamiltonian, $H = \sum_a E_a|a\rangle\langle a|$,

$$\rho = \underbrace{\sum_a \rho_{aa}|a\rangle\langle a|}_{\text{probabilities}} + \underbrace{\sum_{a\neq b} \rho_{ab}|a\rangle\langle b|}_{\text{coherences}}, \quad \text{with } \rho_{ab} = \langle a|\rho|b\rangle = \mathsf{Tr}\left(|a\rangle\langle b|^{\dagger}\rho\right), \tag{24}$$

We put all the components ρ_{ab} into a supervector, and order the components such that the *probabilities* ρ_{aa} lie in the first block of elements labeled with nonitalic subscript "p" and the remaining *coherences* ρ_{ab}, $a \neq b$ form another block labeled "c":

$$\rho_{\mathrm{p}} = \begin{bmatrix} \vdots \\ \rho_{aa} \\ \vdots \end{bmatrix} \quad \text{probabilities}, \qquad \rho_{\mathrm{c}} = \begin{bmatrix} \vdots \\ \rho_{ab} \\ \vdots \end{bmatrix} \quad \text{coherences}. \tag{25}$$

We decompose the supermatrix for L and $W^{(1)}$ into the corresponding blocks labeled by combinations of the non-italic subscripts p and c:

$$\frac{d}{dt}\begin{bmatrix}\rho_{\mathrm{p}} \\ \rho_{\mathrm{c}}\end{bmatrix} = \begin{bmatrix}0 \\ 0\end{bmatrix} = \begin{bmatrix} W^{(1)}_{\mathrm{pp}} & W^{(1)}_{\mathrm{pc}} \\ W^{(1)}_{\mathrm{cp}} & -iL_{\mathrm{cc}} + W^{(1)}_{\mathrm{cc}} \end{bmatrix}\begin{bmatrix}\rho_{\mathrm{p}} \\ \rho_{\mathrm{c}}\end{bmatrix}. \tag{26}$$

The supermatrix elements of the blocks W_{xy} are $\mathsf{Tr}\big[|a\rangle\langle b|^\dagger(W|a'\rangle\langle b'|)\big]$ where $a = b$ ($a \neq b$) for blocks with x=p (c) and $a' = b'$ ($a' \neq b'$) for blocks with y=p (c), respectively.[6] The same holds for the supermatrix of $L = [H, \bullet]_-$ and in Eq. (26) we used that only the cc-block of L is nonzero.[7] Moreover, this block is diagonal and for the following discussion, the key point is to remember is that

$$L_{\text{cc}} \sim E_a - E_b, \tag{27}$$

which has to be compared with the tunnel coupling Γ. To discuss the consistency of the first-order approximation (22) it is useful to formally solve the second of the two coupled stationary equations (26) for the coherences in terms of the probabilities, and then substitute this into the first of the equations (26). This leads to an *effective* equation for the probabilities only:

$$0 = \left(W^{(1)}_{\text{pp}} - W^{(1)}_{\text{pc}} \frac{1}{-iL_{\text{cc}} + W^{(1)}_{\text{cc}}} W^{(1)}_{\text{cp}}\right) \rho_{\text{p}}. \tag{28}$$

This equation has the structure familiar from perturbation theory of the Brillioun-Wigner type: the second term represents the effect on the probabilities of "virtual transitions" to the coherences. We now consider the following cases:

1. *Symmetry*: If by a symmetry or conservation law the off-diagonal supermatrix blocks vanish,

   $$W^{(1)}_{\text{pc}} = 0 = W^{(1)}_{\text{cp}}, \tag{29}$$

   then only the probabilities matter and Eq. (28) reduces to a so-called master equation:

   $$0 = W^{(1)}_{\text{pp}} \rho_{\text{p}}. \tag{30}$$

   For example, the charge conservation for the total system leads to Eq. (30) for simple quantum dot models in which the charge is the only quantum number. In less simple models, the charge conservation still prevents the coherences ρ_{ab} between states with different charge $N_a \neq N_b$ from coupling to any of the probabilities ρ_{aa}. Often also, only spin doublets ($S = 1/2$) appear in simple models and then conservation of the spin component along some physical axis is enough to also decouple the coherences between spin-degenerate states.

2. *Nonsecular coherences*: If the previous does not apply, but instead we have

   $$L_{\text{cc}} \sim E_a - E_b \quad \gg \quad \Gamma \sim W_{\text{cc}}, \tag{31}$$

   then we must still neglect the second term in Eq. (28) and again obtain Eq. (30) in which only the probabilities are relevant, even though there are nonzero coherences. We will see that this so-called secular approximation *breaks down* when going to second order.

[6] To obtain Eq. (26), decompose the operator $W\rho$ (the result of the action of the superoperator) as indicated by Eq. (24) and then insert the expansion Eq. (24) for ρ itself.

[7] Since $|a\rangle$ is the energy basis $\mathsf{Tr}\big[|a\rangle\langle b|^\dagger(L|a'\rangle\langle b'|)\big] = (E_{a'} - E_{b'})\mathsf{Tr}\big[|a\rangle\langle b|^\dagger|a'\rangle\langle b'|\big] = (E_a - E_b)\delta_{aa'}\delta_{bb'}$.

3. *Secular coherences*: If instead the relevant states are quasi-degenerate on the scale of the tunnel coupling,

$$L_{\rm cc} \sim E_a - E_b \quad \lesssim \quad \Gamma \sim W_{\rm cc}, \tag{32}$$

then the coherences ρ_{ab} are called secular. In this case, both the probabilities and the secular coherences are relevant and must be kept. In particular, for $L_{\rm cc} \gg W_{\rm cc}$ Eq. (28) reduces to

$$0 = \left(W^{(1)}_{\rm pp} - W^{(1)}_{\rm pc} \frac{1}{W^{(1)}_{\rm cc}} W^{(1)}_{\rm cp} \right) \rho_{\rm p}. \tag{33}$$

We should either use this effective master equation, or the coupled quantum master equation (26), while keeping *only* the secular coherences and *neglecting* $L_{\rm cc}$.

We now study case (1-2) and case (3) separately in more detail and analyze some examples.

3.1 Probabilities

We first consider the stationary master equation (30). Written out it reads:

$$\frac{d}{dt}\rho_{aa} = 0 = \underbrace{\sum_{b\neq a} W^{(1)}_{aa,bb}\,\rho_{bb}}_{\text{gain } a \leftarrow b} - \underbrace{\left(\sum_{b\neq a} W^{(1)}_{bb,aa}\right)\rho_{aa}}_{\text{loss } a \rightarrow b}, \tag{34}$$

where $W_{aa,bb} := \mathsf{Tr}\big[|a\rangle\langle a|^\dagger (W^{(1)}|b\rangle\langle b|)\big]$ and we used the sumrule $W^{(1)}_{aa,aa} = -\sum_{b\neq a} W^{(1)}_{bb,aa}$, which guarantees probability conservation.[8] This equation has the intuitive interpretation that the probability lost by one state b "flows" to state a and that at stationarity all these flows must balance each other.

Transition rates. Explicit calculation of the supermatrix elements by the diagrammatic expression for $W^{(1)}$ gives for the transition rates

$$W_{aa,bb} = \begin{cases} \displaystyle\sum_{r\sigma'} 2\pi \Big| \sum_{\sigma l} t_{\sigma l \sigma' r} \langle a|d^\dagger_{\sigma l}|b\rangle \Big|^2 f^+_r(E_a - E_b) & N_a > N_b \quad \text{electron enters,} \\ \displaystyle\sum_{r\sigma'} 2\pi \Big| \sum_{\sigma l} t^*_{\sigma l \sigma' r} \langle a|d_{\sigma l}|b\rangle \Big|^2 f^-_r(E_b - E_b) & N_b > N_a \quad \text{electron exits,} \end{cases} \tag{35}$$

where $f^\pm_r(\omega) = (e^{\pm(\omega-\mu^r)/T}+1)^{-1}$ is the Fermi distribution function for electrons / holes, in line with the physical interpretation sketched in Fig. 2.

These expressions coincide with those obtained from very simple, text-book "Golden-Rule" arguments. Indeed, master equations of the type (35) can be derived by much simpler considerations and the real-time machinery may seem to be superfluous. However, as soon as coherences and second-order tunneling effects become important this is no longer true, and, for instance, the often used "Golden-Rule + T-matrix" approach leads to serious problems [2].

Single-electron tunneling and Coulomb blockade. With the rates (35) in hand we can now distinguish the two regimes of central interest both for theory and experiment. To obtain some

[8] This property is preserved term-by-term in *any* order of the perturbation theory for W given by Eq. (18) since G^+_1 always appears on the left in W and by construction $\mathsf{Tr} G^+_1 = 0$ as one can show from (19).

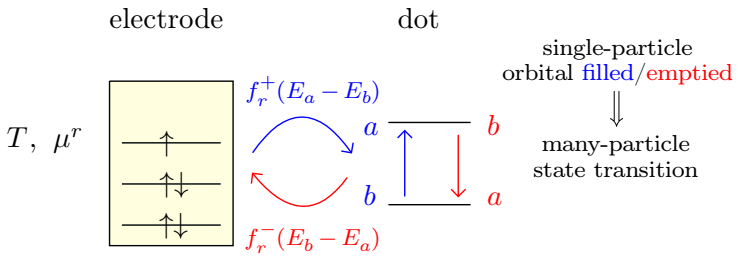


Fig. 2: *First-order tunnel process: an electron tunneling into (blue) or out of (red) the quantum dot induces a transition between two many-body states a and b differing in electron number by $N_a - N_b = \pm 1$. Due to the correlations on the dot, the physically relevant rates are* not *those associated with the filling a particular spin-orbital since that spin-orbital appears in multiple many-body states: for example for the Anderson model [Eq. (43) below] adding a $\uparrow$ electron to the single orbital can induce two many-body state transitions, $|0\rangle \to |\uparrow\rangle$ and $|\downarrow\rangle \to |2\rangle$ with different electrochemical potentials and transition rates.*

general orientation, we first ignore details of the energy spectrum, set $T = 0$ and account only for the dominant Coulomb interaction energy due to capacitive pair interactions of N electrons: $E_N \approx \frac{1}{2}N(N-1)U +$ terms depending on the voltages (also due to capacitive effects) [14]. Here $U = e^2/C$ is the Coulomb charging energy and $C =$ total capacitance. Then, for bias voltages $|\mu^{\rm L} - \mu^{\rm R}| < U$ only two cases are possible, depending on how the gate voltage positions the sequence of energy *differences* $E_N - E_{N-1}$:

- *Single-electron tunneling* (SET) regime: The dot electrochemical potential lies in the transport window, defined by the bias across the electrodes $V_b = \mu^{\rm L} - \mu^{\rm R}$:

$$\mu^{\rm L} > E_{N+1} - E_N > \mu^{\rm R}, \tag{36}$$

  and the inequalities need to be reversed for reverse bias $\mu^{\rm L} < \mu^{\rm R}$. This means that electrons from the left electrode induce $N \to N+1$ and when leaving to the right electrode induce $N+1 \to N$, leading to a directed transport current.

- *Coulomb blockade* (CB) regime: The electrochemical potentials of the electrodes lie between adjacent dot electrochemical potentials

$$E_N - E_{N-1} < \mu^{\rm L}, \mu^{\rm R} < E_{N+1} - E_N. \tag{37}$$

  Now electrons from both electrodes induce $N-1 \to N$ as well as $N \leftarrow N+1$. Therefore, the dot charge is stable and fixed to N. The energy responsible for the resulting blockade of transport is the finite *second* difference $(E_{N+1} - E_N) - (E_N - E_{N-1}) = U$.

These regimes can still meaningfully defined in more general cases when above $x > (<) y$ is taken to mean "larger (smaller) on the scale of thermal (T) and tunneling fluctuations (Γ)" i.e., $x - y \gg (\ll) T, \Gamma$. Also, quite generally the dot energies E_a depend linearly on the applied voltages [14, 15], at least in fairly broad voltage ranges, even for quantum dots of molecular [17] or atomic dimensions [18]. Then, the conditions

$$\mu^{\rm L} \text{ or } \mu^{\rm R} = E_a - E_b \text{ with } N_a = N_b + 1, \tag{38}$$

where E_a and E_b are ground state energies for the respective electron numbers, determine straight lines in the (V_g, V_b) plane [cf. Eq. (38)] that bound the SET regime of these ground

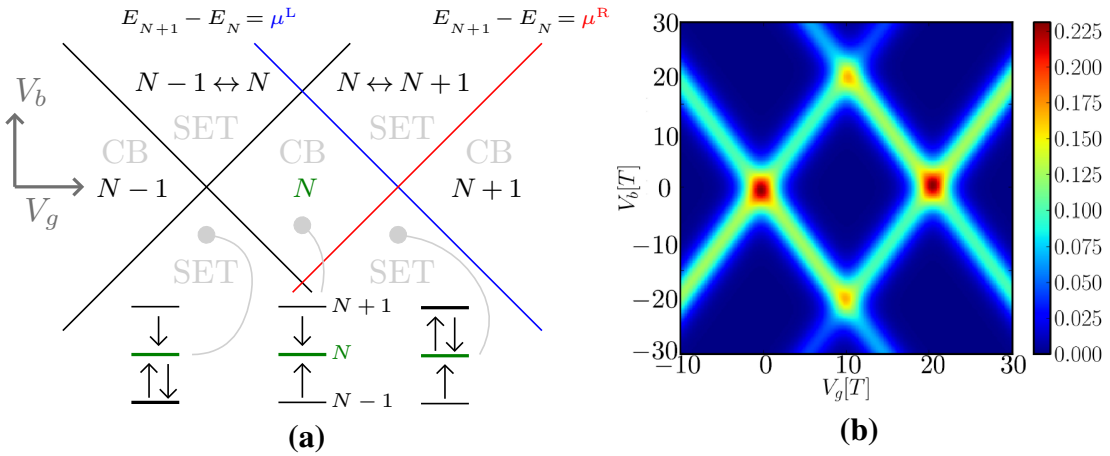


Fig. 3: *First-order probabilities. (a) Sketch of the stability diagram indicating the stable charge states ($N-1$, N, $N+1$) in the respective CB regimes and the charge instabilities ($N-1 \leftrightarrow N$ and $N \leftrightarrow N+1$) in the respective SET regimes. The energy level diagrams (bottom) indicate the direction of energetically allowed transitions in three subsequent regimes. (b) Conductance map, dI/dV plotted vs (V_g, V_b), for a single orbital quantum dot at zero magnetic field. The level $\epsilon = -V_g$ and the bias are given in units of the temperature T, the natural scale of the broadening included in the master equation (30) [cf. Eq. (35)]. The conductance is plotted in units of $\Gamma^R/T = 0.5\Gamma^L/T$, where $\Gamma^r = 2\pi(t^r)^2$ are the spin-independent tunnel rates. The Coulomb charging energy $U = 20T$ determines the width and height of the central "diamond" shaped regime.*

states. By exclusion with other SET regimes for other transitions the various CB regimes are determined. The resulting so called *stability diagram* is sketched in Fig. 3(a). The current changes basically from zero in the CB regime to a value $\sim \Gamma$ in the SET regime. For $T \ll \Gamma$ the resonance line (38) is broadened by Γ due to quantum fluctuations. Restoring all the units, one then estimates the conductance peak values to be $dI/dV_b \sim (e\Gamma)/(h\Gamma/e) \sim e^2/h$, the fundamental quantum of conductance. In the first-order approximation, however, we have small $\Gamma \ll T$ and the broadening is given by T instead. The approximation (22) thus applies for nonlinear conductances which are small on the scale of the quantum conductance, i.e., $dI/dV_b \sim e^2/h(\Gamma/T) \ll e^2/h$.

Example: Single orbital quantum dot - Anderson model. To illustrate this general picture more concretely, consider a quantum dot with a single orbital, giving four many-body states:

$$|0\rangle \qquad N = 0, \text{ singlet}, \tag{39}$$

$$|\sigma\rangle = d_\sigma^\dagger |0\rangle \qquad N = 1, \text{ doublet } \sigma = \uparrow, \downarrow, \tag{40}$$

$$|2\rangle = d_\uparrow^\dagger d_\downarrow^\dagger |0\rangle \qquad N = 2, \text{ singlet}. \tag{41}$$

We introduce the Anderson model, which will be used throughout this chapter:

$$H = \sum_\sigma \epsilon_\sigma N_\sigma + \tfrac{1}{2} N(N-1)U \tag{42}$$

$$= \sum_\sigma \epsilon_\sigma N_\sigma + N_\uparrow N_\downarrow U \qquad = \sum_\sigma \epsilon_\sigma |\sigma\rangle\langle\sigma| + \Big(\sum_\sigma \epsilon_\sigma + U\Big)|2\rangle\langle 2| \tag{43}$$

with $N_\sigma = d_\sigma^\dagger d_\sigma$ and $N = \sum_\sigma N_\sigma$, and we drop the orbital index l in the tunnel matrix elements in Eq. (3). See also lectures B3 by T. Costi and B7 by V. Meden. The dot electrochemical

potentials lie at $E_\sigma - E_0 = \epsilon_\sigma$ and $E_2 - E_\sigma = \epsilon_{\bar\sigma} + U$ where $\bar\sigma = \downarrow, \uparrow$ denotes the opposite of $\sigma = \uparrow, \downarrow$. The dI/dV_b map calculated for this model in the first-order approximation (22) is shown in Fig. 3 and illustrates the general expectation sketched above for the appearance of peaks in the stability diagram. These peaks indicate a step in the current I to a new plateau value as the bias V_b is increased. For simplicity we considered $\epsilon_\uparrow = \epsilon_\downarrow := \epsilon$ (no magnetic field) and the orbital level is assumed to depend on the gate voltage as $\epsilon = -V_g$,[9] and the tunnel coupling is spin independent.

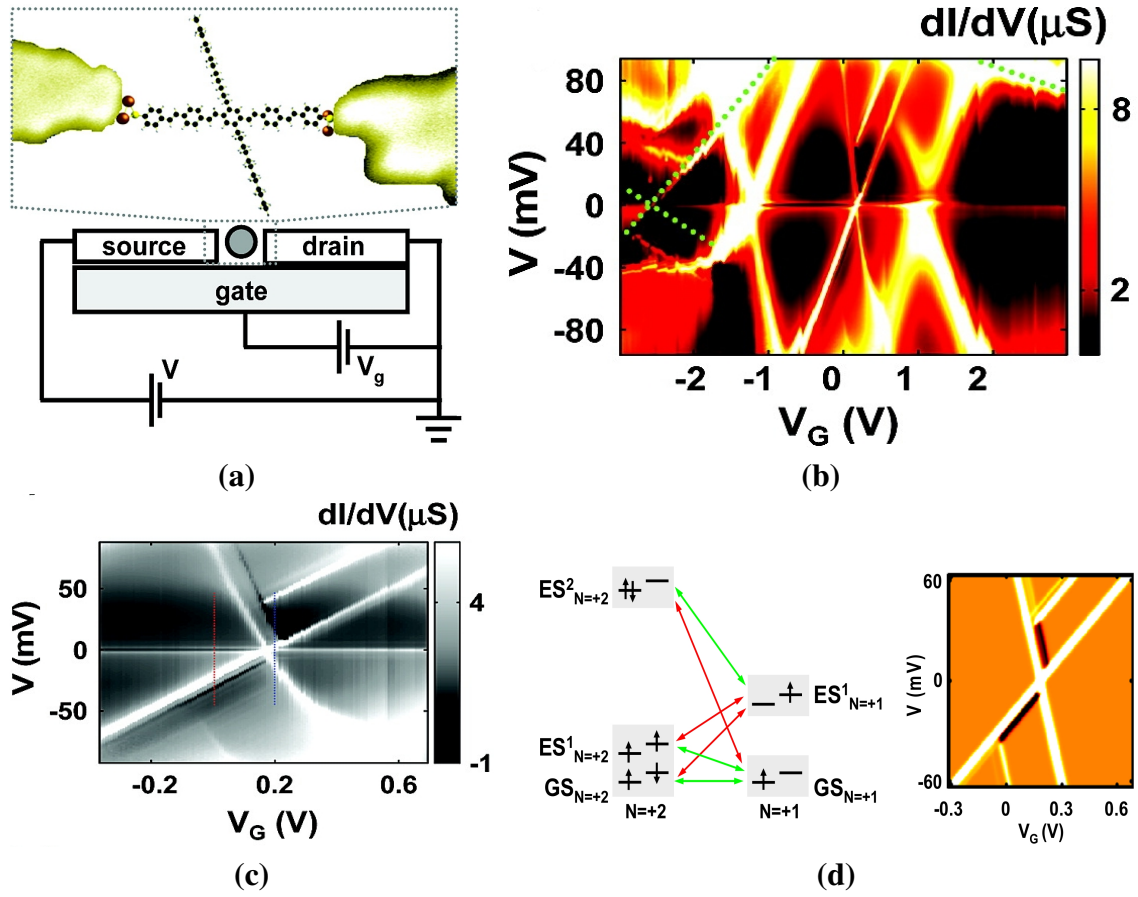


Fig. 4: *OPV5 molecular junction. (a) Sketch of an OPV5 molecule in a junction between gold electrodes (top) and schematic three terminal device setup (bottom). (b) Experimental dI/dV map (stability diagram) and (c) zoomed dI/dV map in the regime where few excitations dominate the transport. (d) First-order perturbation theory: the right panel shows dI/dV (arbitrary units) modeling (c), calculated using the master equation (30) and the model sketched in the left panel (green lines = large tunnel rates, red lines = small tunnel rates). Figure adapted from [19] and its Supplementary Online Material.*

Experiment: OPV5 molecular double dot. In Fig. 4 we show experimental nonlinear conductance plots for a three-terminal quantum dot device made up of an nanometer scale organic OPV5 molecule containing five conjugated rings and two side arms [19]. Remarkably, many of its basic features are qualitatively quite well described in sizeable ranges of the applied voltages

[9] This is by an appropriate choice of the energy origin and ignores capacitive conversion factors [14] for simplicity. The sign is however crucial: positive gate voltage attracts electrons to the dot, lowering the electrochemical potentials $E_N - E_{N-1}$, allowing electrons to "fall in" easier.

by the very simple type of model outlined above. This indicates that due to the strong Coulomb correlation effects on the molecule transport is mostly limited to *single*-electron tunneling, despite the presence of many electrons on both molecule and electrodes. Based on the signatures of transport alone – also involving second-order processes to be discussed in Sec. 4 and the Kondo effect discussed in lectures B3, B5 and B7 – a quite detailed model could be worked out for this particular junction, even though the molecule could not be detected by any other means. See [19] for the details. Here we highlight the feature close to 0.2 V in gate voltage in the bias range of about -60 to 60 meV, shown in Fig. 4(c). In addition to the stability diagram features discussed above in Fig. 3, electronic excitations of the molecule show up as additional lines. Interestingly, the conductance at one of these resonances is *negative*, it is a dI/dV_b dip as function of V_b, not a peak. This indicates that some of the new states that become accessible at higher energy have smaller tunnel rates associated with them. Fig. 4(d) shows that a very crude model, which only makes assumptions on the relative values of these rates (see sketch), can reproduce the features of the data in Fig. 4(c) in some detail by solving Eq. (30). This provides valuable information: the fact that the negative conductance shows up, indicates that the relaxation of the molecular excitations is *slower* than the transport processes. The time scale given by the inverse of the measured current thus provides a lower bound on the internal relaxation time. Other detailed comparisons of SET theory with molecular transport experiments can be found, e.g., in [20, 21, 22, 23, 24, 25, 26, 27].

3.2 Secular coherences

We now turn to cases where there is no symmetry such that $W^{(1)}_{\rm pc} = 0 = W^{(1)}_{\rm cp}$ and there are secular coherences, in particular we consider

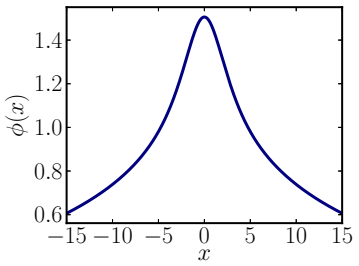
$$L_{\rm cc} \sim E_a - E_b \quad \ll \quad \Gamma \sim W_{\rm cc}. \tag{44}$$

Then the master equation description (30) in terms of probabilities alone breaks down and instead a kinetic equation or quantum master equation is required in which $L_{\rm cc}$ must be neglected:

$$\frac{d}{dt}\begin{bmatrix}\rho_{\rm p}\\ \rho_{\rm c}\end{bmatrix} = 0 = \begin{bmatrix}W^{(1)}_{\rm pp} & W^{(1)}_{\rm pc}\\ W^{(1)}_{\rm cp} & W^{(1)}_{\rm cc}\end{bmatrix}\begin{bmatrix}\rho_{\rm p}\\ \rho_{\rm c}\end{bmatrix}. \tag{45}$$

We thus need to evaluate additional blocks $W^{(1)}_{\rm pc}$, $W^{(1)}_{\rm cp}$, and $W^{(1)}_{\rm cc}$ of the generalized rate supermatrix. When doing this [1], a new type of function arises that has a very different profile as compared with the Fermi function $f^+(x) = (e^x + 1)^{-1}$ appearing in Eq. (35):

$$\begin{aligned} W_{\rm cc} \sim \phi(x) &= \frac{1}{\pi}\int_{-D/T}^{D/T} d\omega \frac{f^+(\omega)}{x-\omega} \\ &= \frac{1}{\pi}\Big[-\,{\rm Re}\,\Psi\left(\frac{1}{2} + i\frac{x}{2\pi}\right) + \ln\left(\frac{D}{2\pi}\right)\Big] \end{aligned} \tag{46}$$



We refer to $\phi(x)$ as the *level shift function*. As the plot for large band width $D = 100T$ shows, it is maximal close to resonance where $x = (E_a - E_b - \mu^r)/T \ll 1$, i.e., where the repulsion experienced by the dot levels from the filled and empty levels of the electrode is the strongest. To get an idea of what this physically corresponds to we now consider a concrete example.

Example: Quantum dot spin-valve. We again consider the Anderson model (43), but we now allow the density of states for spin up and down in the electrodes to be different as sketched in Fig. 5. Not sketched in this figure, but crucial here, is that the spin polarizations in the two electrodes point in different directions given by two vectors $\mathbf{n}^{\mathrm{L}}$ and $\mathbf{n}^{\mathrm{R}}$.[10] We now have to be very careful with the spin-coherences: given some spin-quantization axis on the dot we can decompose the density operator as

$$\rho = \rho_{00}|0\rangle\langle 0| + \sum_{\sigma} \rho_{\sigma\sigma}|\sigma\rangle\langle\sigma| + \rho_{22}|2\rangle\langle 2| + \rho_{\uparrow\downarrow}|\uparrow\rangle\langle\downarrow| + \rho_{\downarrow\uparrow}|\downarrow\rangle\langle\uparrow| + \dots . \tag{47}$$

where the omitted terms involve coherences that decouple from the probabilities by charge conservation. In Eq. (45) we then concretely have (note that $\rho_{\downarrow,\uparrow} = \rho_{\uparrow,\downarrow}*$ since $\rho^\dagger = \rho$)

$$\rho_{\mathrm{p}} = \begin{bmatrix} \rho_{00} \\ \rho_{\uparrow\uparrow} \\ \rho_{\downarrow\downarrow} \\ \rho_{22} \end{bmatrix} \quad \text{probabilities,} \qquad \rho_{\mathrm{c}} = \begin{bmatrix} \rho_{\uparrow,\downarrow} \\ \rho_{\downarrow,\uparrow} \end{bmatrix} \quad \text{coherences.} \tag{48}$$

This decomposition depends on the choice of quantization axis. If the spin-polarization axes of the ferromagnets are parallel one can chose to quantize the spin on the dot along this *physical* axis. In this basis, one can show that the rate supermatrix blocks coupling the coherences to the probabilities are exactly zero, i.e., $W^{(1)}_{\mathrm{pc}} = 0 = W^{(1)}_{\mathrm{cp}}$, due to spin-rotation symmetry with respect to this axis. However, for noncollinear polarizations this argument breaks down: by no choice of quantization axis on the dot can such a decoupling be realized. Physically this is clear: the axial symmetry of the problem has been broken. The coherences have to be kept to get the correct physical description. This also leaves one free to choose the axis in a clever way [29] or derive kinetic equations in a form explicitly independent of the choice of coordinate system [28, 30].

To see more clearly what the coherences physically describe we make another decomposition according to the transformation behavior of the components of the density operator under spin-rotations:

$$P = \begin{bmatrix} \rho_{00} \\ \rho_{\uparrow\uparrow} + \rho_{\downarrow\downarrow} \\ \rho_{22} \end{bmatrix} \quad \text{scalar probabilities,} \quad S = \begin{bmatrix} \mathrm{Re}\,\rho_{\uparrow,\downarrow} \\ \mathrm{Im}\,\rho_{\uparrow,\downarrow} \\ \frac{1}{2}\left(\rho_{\uparrow\uparrow} - \rho_{\downarrow\downarrow}\right) \end{bmatrix} \quad \text{vector components,} \tag{49}$$

where we used $\rho^*_{\uparrow,\downarrow} = \rho_{\uparrow,\downarrow}$. The kinetic equation (45) can be rewritten as

$$\frac{d}{dt}\begin{bmatrix} P \\ S \end{bmatrix} = 0 = \begin{bmatrix} W^{(1)}_{\mathrm{PP}} & W^{(1)}_{\mathrm{PS}} \\ W^{(1)}_{\mathrm{SP}} & W^{(1)}_{\mathrm{SS}} \end{bmatrix} \begin{bmatrix} P \\ S \end{bmatrix}. \tag{50}$$

There is thus a coupling of occupation (charge) and spin dynamics. One shows along similar lines that the experimentally measurable charge current couples to the spin accumulation S as well. The explicit calculation of the supermatrix blocks $W^{(1)}_{\mathrm{PS}}$, $W^{(1)}_{\mathrm{SP}}$ and $W^{(1)}_{\mathrm{SS}}$ was first done by Braun et. al. [29]: besides vectorial rates generating and attenuating the spin accumulation ($W^{(1)}_{\mathrm{PS}}$, $W^{(1)}_{\mathrm{SP}}$), one obtains the components of a rank 2 tensor ($W^{(1)}_{\mathrm{SS}}$) describing spin relaxation

[10] This information goes into the spin dependence of the tunnel matrix elements $t_{\sigma l\sigma' r}$ in Eq. (3), see [28].

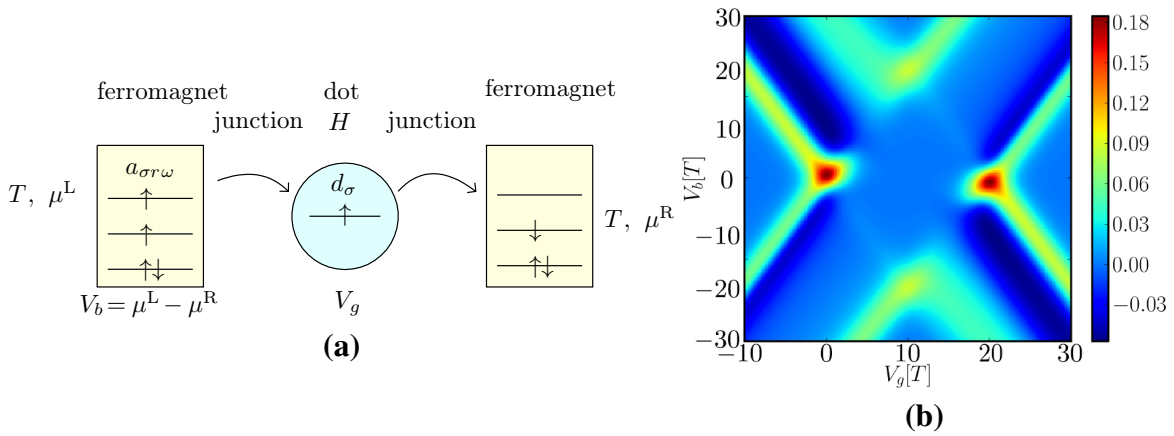


Fig. 5: *First-order secular coherences: quantum-dot spin valve. (a) Single orbital quantum dot coupled to ferromagnets. (b) dI/dV map calculated using the first-order kinetic / quantum master equation (45), or equivalently, Eq. (50), which includes the effect of the secular coherences. The polarization vectors of the ferromagnets are perpendicular and the polarization of the DOS is taken to be 99% as in Fig. 3(b) dI/dV is plotted in units of $\Gamma^R/T = 0.5\Gamma^L/T$ and all further parameters are the same.*

(symmetric part) and *spin precession* (antisymmetric part). The latter is equivalent to an *exchange field* pseudovector

$$\mathbf{B} = \sum_r \Gamma^r \mathbf{n}^r \left[-\phi(\epsilon - \mu^r) + \phi(\epsilon + U - \mu^r)\right], \tag{51}$$

which now illustrates how the level-renormalization function (46) enters the problem. The correct description of the precession effect requires the inclusion of the secular coherences. Physically, this renormalization is due to the quantum fluctuations of an electron by which it "probes" the spin-polarization in the electrodes and transfers it to the strongly correlated dot through Eq. (51). This is a correlations effect: for $U = 0$ Eq. (51) gives $\mathbf{B} = \mathbf{0}$. This may have a big impact on the conductance, which is plotted in Fig. 5: as compared to the result for nonmagnetic electrodes in Fig. 3(b), there is a strong suppression of some of the SET resonances and also negative dI/dV values occur, indicating a significant current suppression. In fact, the spin precession is responsible for the current *not* being suppressed *everywhere* [29]: injection of electrons from the source tends to induce a spin accumulation that is noncollinear with the drain electrode. This prevents an electron from leaving the dot by the spin-valve effect. However, the precession of the spin accumulation on time scales $|\mathbf{B}|^{-1}$ allows the electron waiting on the dot for a typical time Γ^{-1} – which is comparable [cf. Eq. (51)] – to precess its spin into the direction of the drain electrode, and thereby escape. Since the two terms in the exchange field (51) tend to cancel out between the SET resonances, one expects that the spin-valve suppression can take place there. This leads to the broad band (with negative slope) of negative dI/dV_b values followed by a parallel broad band of positive dI/dV_b in Fig. 5(b).

The coherence effects that we studied in the above, very basic model continue to be discussed [29, 31, 32, 33, 34, 35]. For instance, quantum dots with higher spin values have also been studied. These can exhibit other types of accumulations than the spin vector $\mathbf{S}$ and other spintronic exchange fields than the vector (51) which are associated with higher spin multipoles in particular the spin *quadrupole* tensor [28, 36, 30, 37]. Also, effects related to secular coherences have been predicted for devices with *orbital* (rather than spin) degrees of freedom, such

as double quantum dots [38] and molecular devices [39, 40, 41, 42, 43, 44, 45], and vibrational degrees of freedom [46].

4 Second-order tunneling effects

So far we have focused on effects close to the SET regime. Now we want to describe the Coulomb blockade regime, where experimentally often important excitations can be detected with increased accuracy [47]. One first needs to extend the calculation of the rate supermatrix to the second order as outlined in Sec. 2.2:

$$W \approx W^{(1)} + W^{(2)}, \tag{52}$$

where $W^{(n)} \propto \Gamma^n$ for $n = 1, 2$. As in the case of the first-order approximation, this is not the end of the story yet: We have to consider the consistency of the stationary equation

$$\frac{d}{dt}\begin{bmatrix}\rho_{\rm p}\\ \rho_{\rm c}\end{bmatrix} = 0 = \begin{bmatrix} W^{(1)}_{\rm pp} + W^{(2)}_{\rm pp} & W^{(1)}_{\rm pc} + W^{(2)}_{\rm pc} \\ W^{(1)}_{\rm cp} + W^{(2)}_{\rm cp} & -iL_{\rm cc} + W^{(1)}_{\rm cc} + W^{(2)}_{\rm cc}\end{bmatrix}\begin{bmatrix}\rho_{\rm p}\\ \rho_{\rm c}\end{bmatrix}. \tag{53}$$

Here, we now consider two important cases:

1. *Symmetry*: If a symmetry makes the blocks coupling to the coherences vanish,

   $$W^{(1)}_{\rm pc} + W^{(2)}_{\rm pc} = 0 = W^{(1)}_{\rm cp} + W^{(2)}_{\rm cp}, \tag{54}$$

   then again only probabilities are relevant and we have a second-order master equation:

   $$0 = \left(W^{(1)}_{\rm pp} + W^{(2)}_{\rm pp}\right)\rho_{\rm p}. \tag{55}$$

2. *Nonsecular coherences*: In contrast to the first-order case, if there is no such symmetry coherences can *not* be neglected, *even* if they are nonsecular, the case we consider now:

   $$L_{\rm cc} \sim E_a - E_b \quad \gg \quad \Gamma \sim W^{(1)}_{\rm cc}. \tag{56}$$

   In this limit the nonsecular coherences that can (and should) always be neglected in first order [cf. Sec. 3] reappear: eliminating the coherences from Eq. (53) as before, we obtain the effective equation

   $$0 = \left(W^{(1)}_{\rm pp} + W^{(2)}_{\rm pp} - W^{(1)}_{\rm pc}\frac{1}{-iL_{\rm cc}}W^{(1)}_{\rm cp}\right)\rho_{\rm p}. \tag{57}$$

   The second contributions also scales as Γ^2 and is of equal importance as the second-order elements of $W^{(2)}_{\rm pp}$ (see discussion of Fig. 7 below). In fact, it can be shown that this term can be elegantly incorporated in the diagrammatic framework [2] for the calculation of second-order term $W^{(2)}_{\rm pp}$.[11]

Thus the secular approximation, often applied in first order, always breaks down in second order, unless the model exhibits a sufficient amount of symmetries that prohibits this. The effect of *secular coherences* in second-order perturbation theory will not be discussed here [2, 48]. The main point to remember is that once the rate supermatrix W includes higher order tunnel processes, the solution of the stationary state equation (15) requires care. We now consider case (1) and (2) more closely and illustrate them with some interesting applications.

[11] The second term can also be related to the projection onto probabilities in the Nakayima-Zwanzig approach as discussed in [16].

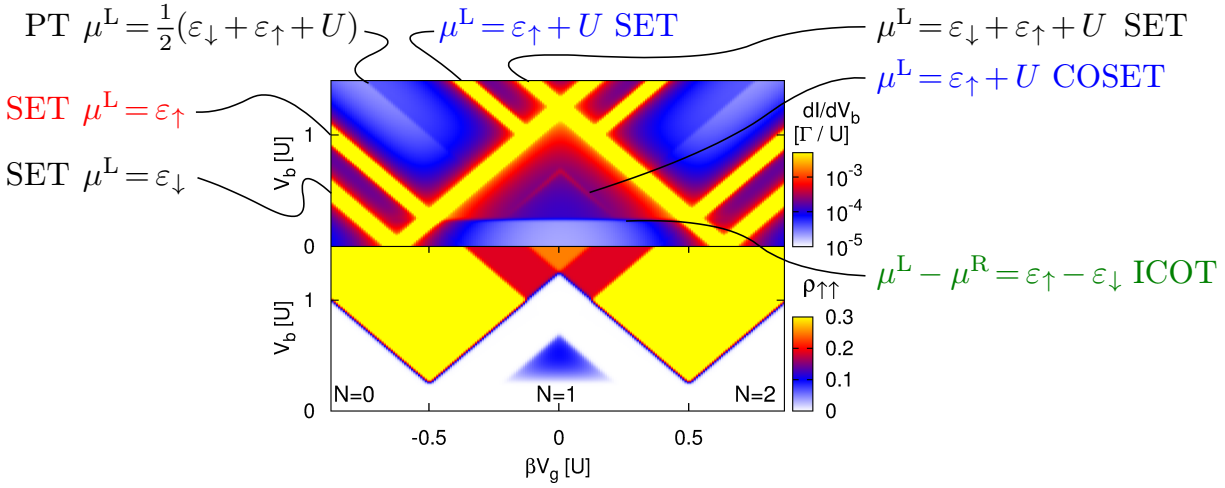


Fig. 6: *Second-order probabilities. (upper panel): dI/dV map for a single orbital quantum dot as a function of bias voltage V_b and gate voltage βV_g. Here β is the gate coupling factor and level energy origin is set to the center of the figure. (lower panel): corresponding map of occupation of excited spin-state, $\rho_{\uparrow\uparrow}$. The spin degeneracy is lifted by an applied magnetic field: $B = \varepsilon_\uparrow - \varepsilon_\downarrow = 50T$ in units of the electron temperature T. The dot is symmetrically coupled to the left and right electrodes: $\Gamma^L = \Gamma^R := \Gamma = 10^{-2}T = 5 \times 10^{-5}U$. For the description see the text. Figure adapted from [7]. .*

4.1 Probabilities

We first consider the case (54) where symmetries (charge and spin conservation) reduce the problem to the second order to a master equation (55) and the coherences play no role.

Example: Inelastic cotunneling, assisted SET, and electron pair-tunneling. As in the case of first order, we first classify the different resonances that can appear when solving Eq. (55). We again use the single orbital Anderson model (43) with spin-independent tunneling and electrodes, but with a magnetic field B (in energy units), such that $\epsilon_\uparrow - \epsilon_\downarrow = B$. In Fig. 6 we plot the dI/dV_b map obtained by solving (55) [1, 49].

For orientation, we first identify the first-order effects. Due to the magnetic field B, the state $|\downarrow\rangle$ is the ground state for $N = 1$ electrons and $|\uparrow\rangle$ an excitation for the *same* electron number of the dot. The first order $N = 0 \leftrightarrow 1$ SET resonances [cf. Eq. (38)] are now Zeeman-split into pairs

$$\mu^L \text{ or } \mu^R = E_\downarrow - E_0 = \epsilon_\downarrow \qquad \text{(ground to ground state)}, \tag{58}$$

$$\mu^L \text{ or } \mu^R = E_\uparrow - E_0 = \epsilon_\uparrow \qquad \text{(ground to excited state)}. \tag{59}$$

Similarly, there are Zeeman split $N = 1 \leftrightarrow 2$ SET lines

$$\mu^L \text{ or } \mu^R = E_2 - E_\downarrow = \epsilon_\uparrow + U \qquad \text{(ground to ground state)}, \tag{60}$$

$$\mu^L \text{ or } \mu^R = E_2 - E_\uparrow = \epsilon_\downarrow + U \qquad \text{(excited to ground state)}. \tag{61}$$

The four resonances for the case μ^L in (58)-(61) are the lines with negative slope and are identified in Fig. 6. For the case μ^R these are the corresponding lines with positive slope which require no further discussion. Importantly, the resonances involving the excited state $|\uparrow\rangle$ [Eq. (59) and (61)] *terminate* when they hit the boundary of the Coulomb blockade regime if we restrict the

calculation to $W^{(1)} \sim \Gamma$. However, in the second-order approximation (52), several new features appear in Fig. 6, which we now address.

- *Inelastic cotunneling* (ICOT). In Fig. 6 the SET excitation resonance $\mu^{\rm L} = \epsilon_\uparrow$ connects[12] to a *horizontal* line determined for $\mu^{\rm L} > \mu^{\rm R}$ by

$$\mu^{\rm L} - \mu^{\rm R} = E_\uparrow - E_\downarrow = \epsilon_\uparrow - \epsilon_\downarrow = B. \tag{62}$$

  This gate-voltage independent resonance has a simple physical interpretation: from this bias voltage on, an electron with incoming energy $\mu^{\rm L}$ can exit with energy $\mu^{\rm R}$ while flipping its spin in a single, coherent second-order process called inelastic cotunneling (ICOT) or inelastic electron tunneling (IETS) [50]. The energy lost by the tunneling electron is then sufficient to excite the dot by $E_\uparrow - E_\downarrow = \epsilon_\uparrow - \epsilon_\downarrow = B$, flipping the spin on the dot, without changing the dot electron number. Below this voltage the dot cannot be excited by any process of first or second order, as can be seen in the Fig. 6: the resonance line (59) (marked red) does not continue below the ICOT threshold (62) (marked green). If it would not, the dot relaxes without being excited which is unphysical. This will be important in Sec. 4.2.

- *Cotunneling-assisted single-electron tunneling* (COSET). By contrast, the higher SET excitation resonance (61) (marked blue) in Fig. 6, involving the same excited state, does not connect to a horizontal line. Instead, this resonance can be seen to weakly continue into the Coulomb blockade regime, before terminating at the horizontal ICOT line (62) as discussed above. Since above the ICOT line the dot is excited excited to $|\uparrow\rangle$ this feature signals the *relaxation* to the $N = 1$ ground state by two, *first-order* single-electron processes $|\uparrow\rangle \to |N\rangle$ and $|N\rangle \to |\downarrow\rangle$ where $N = 0$ or 2. At these so-called cotunneling-assisted SET (COSET) resonances this relaxation path is "switched" on. This is clearly seen in the lower panel of Fig. 6 where the occupation probability $\rho_{\uparrow\uparrow}$ is plotted: above the ICOT threshold the dot can be excited, but once a COSET line is crossed the excitation is quenched since the first-order relaxation is much faster than the second-order ICOT excitation. Experimentally, the above described behavior of the SET resonance lines (59) and (61) provides an important consistency check on the assignment of transitions.[13]

- *Electron pair tunneling* (PT). A complete calculation of all the contributions to $W^{(2)}$ [1, 49] also predicts transitions between states that differ by a pair of electrons $|2\rangle$ and $|0\rangle$. Due to the quantum coherence of this pair tunneling process the energy that one electron lacks can "borrowed" by the other [49] to overcome the Coulomb repulsion energy U. This is expressed by the resonance condition

$$2\mu^{\rm L} \text{ or } 2\mu^{\rm R} = E_2 - E_0 = \epsilon_\downarrow + \epsilon_\uparrow + U. \tag{63}$$

[12] For $\mu^{\rm L} \gtrless \mu^{\rm R}$, at the point where the line $\mu^{\rm L}$ or $\mu^{\rm R} = E_\uparrow - E_0$ (ground to excited state) hits the boundary determined by $\mu^{\rm R}$ or $\mu^{\rm L} = E_\downarrow - E_0$ (ground to ground state), subtracting the two conditions gives $|\mu^{\rm L} - \mu^{\rm R}| = E_\uparrow - E_\downarrow$, i.e., the condition for ICOT is satisfied. The ICOT horizontal line thus goes through this point as well.

[13] When a lower SET excitation (59) connects to the ICOT feature while a corresponding higher SET excitation (61) does not, but instead continues as a COSET feature, then these two features are a signature of *one and the same* excitation of the quantum dot. If they don not match in this way, then in order to describe the experiment one needs to add more excitations to the model. ICOT lines that do not connect to a SET excitation are suspect of belonging to an independent object in the junction transporting current in parallel.

Rewritten as $\mu^{\rm L}$ or $\mu^{\rm R} = \frac{1}{2}(\epsilon_\downarrow + \epsilon_\uparrow + U)$ this shows that the pair-tunneling resonance, marked in Fig. 6 lies just half-way between two single-electron resonances (58) and (60). By this coherent pair tunnel process the two electron state can thus be reached much earlier than by two, single-electron processes. Similar to the ICOT this leads to step in dI/dV_b, i.e., a change in the color, and not a peak. However, in contrast to the ICOT condition Eq. (62), the PT resonance position depends on the gate voltage. The appearance of this pair tunneling resonance, separated from the SET lines, is a real *correlation effect*, since for $U = 0$ all these resonances all coincide.

Numerous experiments have observed the ICOT and COSET features and early works are [47, 51]. One reason for this is that at low temperature nonperturbative effects enhance the ICOT features, adding a peak on top of the step[14] that we found perturbatively. This peak signals the onset of the nonequilibrium Kondo effect [52], which can be described within the real-time formalism as well [53] by extending it with the RG approach presented in lecture B5 by H. Schoeller, see also lectures B3 by T. Costi and B7 by V. Meden. The pair tunneling has so far not been reported in experiments, perhaps for the following reasons. First, the effect has not been noted until recently [49], probably because, naively speaking, "adding two electrons to a quantum dot should be forbidden by Coulomb interactions" and no one bothered to calculate *all* the second-order contributions that disprove this incorrect intuition. Without theoretical guidance, there has been no experimental effort to look for the effect. Additionally, (unpublished) calculations show that additional electronic levels in a quantum dot lead to a cancellation of effects that can suppress pair tunneling resonances, indicating that special conditions may be required to observe it. Finally, unlike the ICOT, the pair tunneling resonance does not seem to undergo a nonperturbative low-temperature enhancement similar to the ICOT as a recent real-time RG study [3] accounting for charge-fluctuations seems to indicate (no Kondo enhancement was included so far). Pair tunneling for effectively attractive interactions has been studied in [54].

4.2 Nonsecular coherences

We now turn to the case (56) where in second order the *first-order nonsecular coherences* become relevant. Quite generally, this is the case for quantum dots with complex spectra where the states have many quantum numbers and few symmetries such that the blocks $W^{(1)}_{\rm pc}$ and $W^{(1)}_{\rm cp}$ in Eq. (57) do not vanish. We thus need to solve the stationary kinetic or quantum master equation (53), neglecting $W^{(2)}_{\rm pc}$, $W^{(2)}_{\rm cp}$, and $W^{(1)}_{\rm cc} + W^{(2)}_{\rm cc}$:

$$\frac{d}{dt}\begin{bmatrix}\rho_{\rm p}\\ \rho_{\rm c}\end{bmatrix} = 0 = \begin{bmatrix} W^{(1)}_{\rm pp} + W^{(2)}_{\rm pp} & W^{(1)}_{\rm pc} \\ W^{(1)}_{\rm cp} & -iL_{\rm cc}\end{bmatrix}\begin{bmatrix}\rho_{\rm p}\\ \rho_{\rm c}\end{bmatrix}. \tag{64}$$

Example: Molecular vibration. An interesting example where this becomes important is obtained when extending the single orbital Anderson model to include a vibrational mode with frequency ω:

$$H = \epsilon\sum_\sigma N_\sigma + N_\uparrow N_\downarrow U + \tfrac{1}{2}\omega(p^2+q^2) + \sqrt{2}\omega\lambda q\sum_\sigma N_\sigma. \tag{65}$$

with $N_\sigma = d^\dagger_\sigma d_\sigma$ and $q = (b^\dagger + b)/\sqrt{2}$ is the vibrational coordinate normalized to the zero-point motion and $p = i(b^\dagger - b)/\sqrt{2}$ the conjugate momentum. Due to the linear Holstein

[14] In there is a small overshoot already in pertubation theory, see discussion of Fig. 10(b).

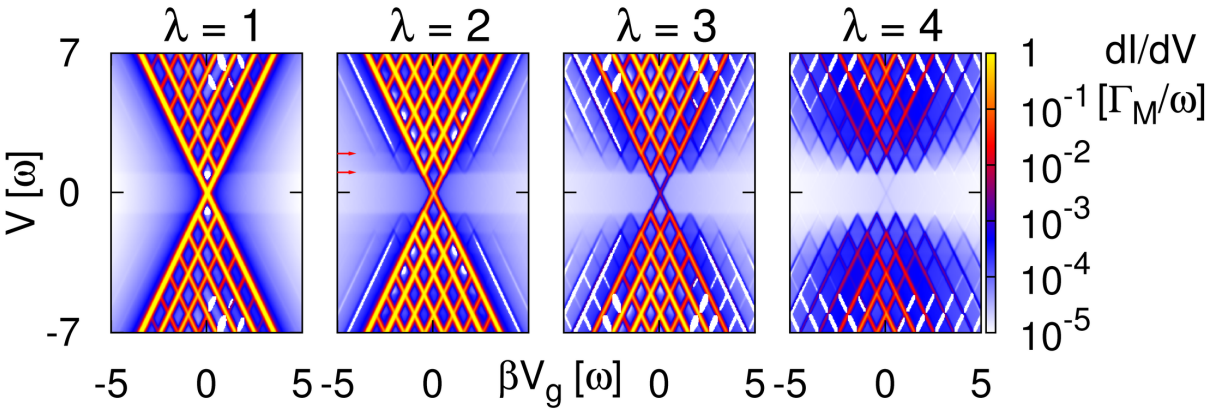


Fig. 7: *Nonsecular coherences and vibration assisted tunneling.* dI/dV *map (*$\beta =$ *gate coupling factor) for a molecular level coupled to a vibrational mode with frequency* ω *and strong local interaction* $U \gg T, V, \lambda$ *[1]. The red arrows in the* $\lambda = 2$ *figure indicate* absorption *of one and two vibrational quanta. The white lines / regions correspond to negative* dI/dV*, which cannot be plotted due to the logarithmic color scale. In all plots we have chosen* $\omega = 40T = 10^4\Gamma$ *and conduction bandwidth* $D = 250\omega$*. The conductance is scaled to the maximal sequential tunneling rate* Γ_M*, i.e.,* Γ *times the maximal Franck-Condon overlap factor (which is less than 1). Figure adapted from [7].*

coupling λ the vibrational mode is displaced when electrons occupy the molecular quantum dot. By the Franck-Condon effect the first-order transition rates become suppressed with increasing λ [55, 56, 57, 58, 59] and second-order processes start to take over [60], *even in the SET regime*. For the results shown in Fig. 7 one finds [1] that the correction from the nonsecular coherences, the second term in Eq. (57), is crucial: when leaving it out "by hand" the COSET excitations from the oscillator continue below the horizontal ICOT line $\mu^{\rm L} - \mu^{\rm R} = V_b = \omega$ (not shown). In such an erroneous calculation the oscillator thus relaxes without being excited! When properly including this correction term this spurious effect is canceled, demonstrating the equal importance as compared to the first term in Eq. (57). Interestingly, Fig. 7 illustrates that in such complex quantum dots the COSET features may lead to negative dI/dV_b in the Coulomb blockade regime if there is little intrinsic relaxation of the vibrational excitations (i.e., other than that provided by transport processes). The current in fact shows a *peak* there, which is very uncommon.

Experiment: Carbon-nanotube electron-vibrational COSET. An interesting experimental example of a complex quantum dot for which Eq. (64) is crucial was realized experimentally [61] in a carbon nanotube with both quantized vibrational *and* electronic excitations. Here features were observed that indicate an interplay between ICOT leading to *electronic* excitation ("heating" via the electronic system) and subsequent SET involving *vibrational* transitions ("cooling" the molecule). The experimental data and the theoretical modeling are shown in Fig. 8(a) and (b), respectively. In Fig. 8(b) the arrows in the CB regime indicate the enhancement of COSET beyond the negative bias (~ -4 meV) where the electronic excited state becomes populated by ICOT.[15] Purely electronic COSET was observed in [51].

[15]Note that in Fig. 8(b) the appearance of COSET features below the *electronic* ICOT threshold at $V_{SD} \approx -4$ meV is not at variance with the above discussion of Fig. 7. There is no COSET below the *vibrational* threshold $V_{SD} = \omega = 810\,\mu$eV, which is, however, not visible either due to smallness of the relevant rates.

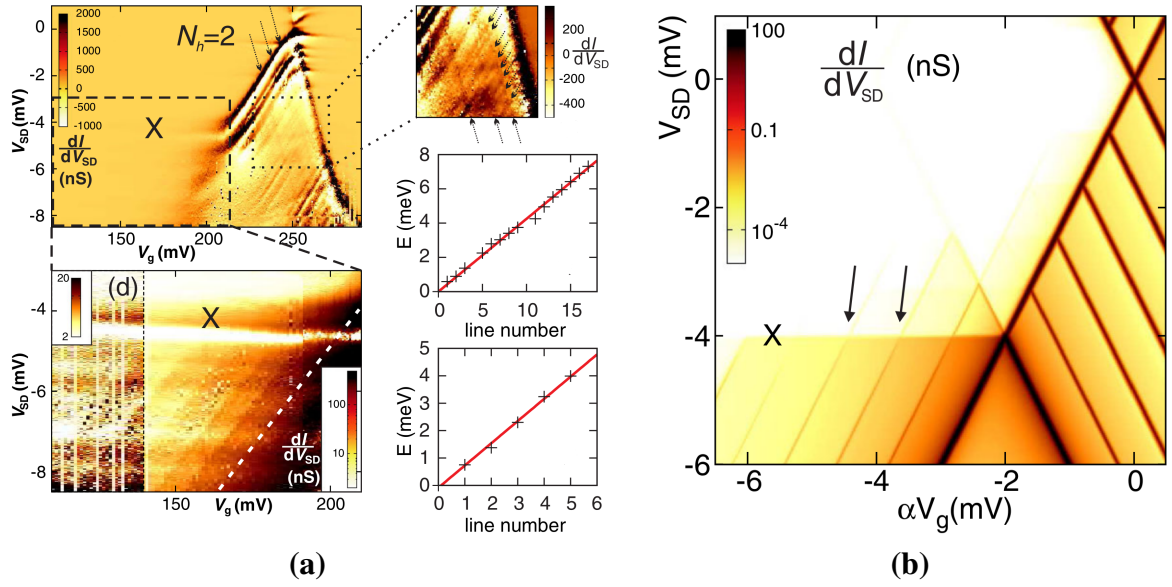


Fig. 8: *Nonsecular coherences and electron-vibrational COSET in a carbon nanotube. (a) Experiment: (top left) dI/dV_{SD} map in the region where the nanotube is charged with $N_h = 1$ or 2 holes. (top right) Detail zoom of the dI/dV_{SD} map in the $N_h = 1 \leftrightarrow 2$ SET regime, black arrows indicating the vibrational excitations. (bottom left) Detail zoom of the dI/dV_{SD} map in the $N_h = 2$ CB region (logarithmic color scale), the white dashed line marking the extrapolated edge of the SET region. The plots on the right show the SET excitation energies versus the number of the excitation line, confirming the harmonic nature of the spectrum. (b) Second-order perturbation theory: dI/dV_{SD} map calculated using the quantum master equation (64) for a simple quantum dot model with equidistant excitations fitted to the observed vibrational level spacing $\omega = 810$ μeV, experimental temperature $T = 8.6\mu$ eV and V_g is scaled with the gate conversion factor $\alpha = C_g/C$. The tunnel couplings were chosen smaller than in the experiment to ensure a well-behaved perturbation expansion. Figure adapted from [61].*

Experiment: Fe_4 single-molecule magnet and spin-tunneling ICOT. As a final example, we consider a quantum dot formed by a molecule that has a large spin $S = 5$ in one of its charge states, a so called single-molecule magnet [62]. Its structure and the experimental junction setup are sketched in Fig. 9. This class of molecules exhibits a very interesting excitation spectrum, which (in a fixed charge state) is well-described by a "giant spin" Hamiltonian

$$H = -DS_z^2 + E(S_x^2 - S_y^2) + \mathbf{B} \cdot \mathbf{S}, \tag{66}$$

with $D > 0$. The microscopic origin of the quadratic terms in this model is discussed in lectures A 10 by G. Bihlmayer on *Relativistic effects* and lecture C4 by S. Blügel on *Complex magnetism*. Here $\mathbf{B}$ is the external magnetic field coupling to the spin. In usual situations, the D-term often dominates over the transverse E-term and raises a barrier opposing the reversal of the spin, e.g., from $|M = S\rangle$ to $|M = -S\rangle$. The transverse term causes quantum spin-tunneling through the D-barrier and has been intensely investigated in the field of molecular magnetism and molecular spintronics. [63] However, the ratio of E/D depends strongly on the geometry and environment of the molecule, and may be quite different from the bulk value when putting such a molecule in a junction. Therefore, when three-terminal experiments were performed on such molecules, [21, 22, 62] the question arose what kind of features the D- and E-term would cause in the nonequilibrium transport [64, 65, 66, 67], in particular in the Coulomb blockade

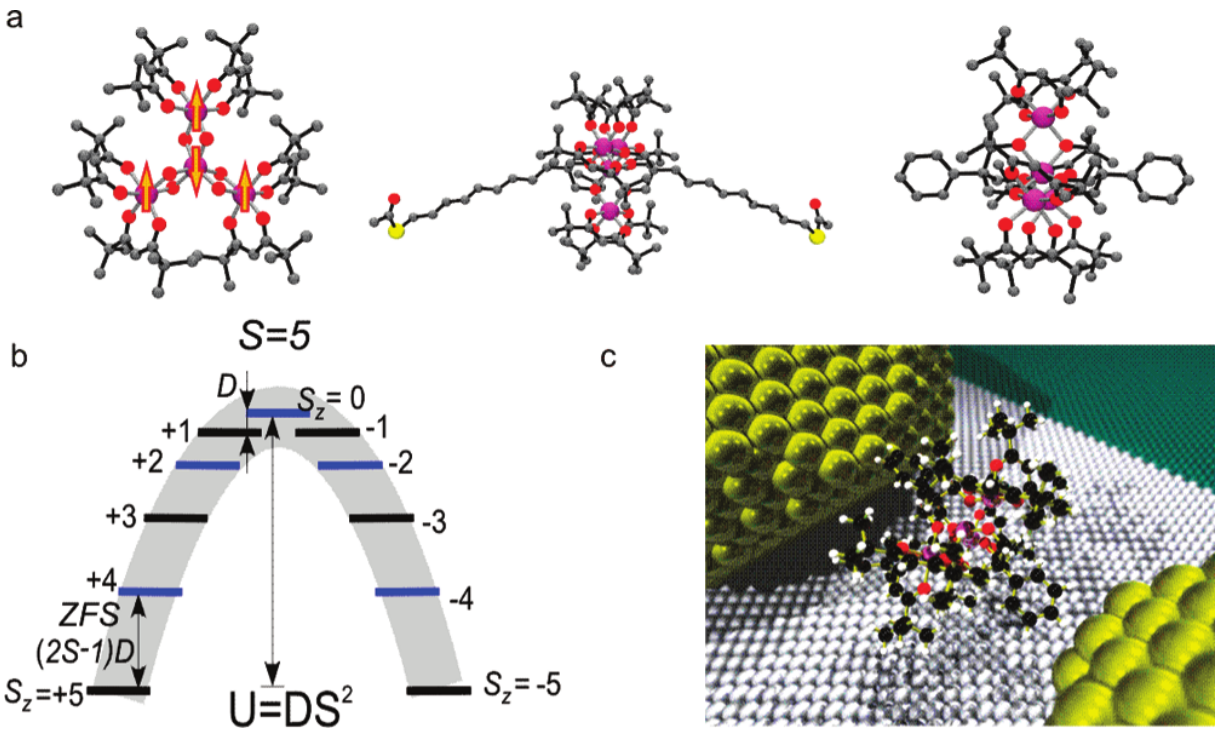


Fig. 9: *(a) Fe_4 single molecule magnet: (left) Fe_4 core with four $S = 5/2$ Fe ions, which due to the antiferromagnetic coupling to the center give a $S = 5$ ground state. (middle, right) Fe_4 with two different linkers groups used for junction formation. (b) Magnetic anisotropy barrier (bottom) $S = 5$ ground state excitation spectrum (left) and sketch of junction (right)*

regime. Indeed, in the experiment [62] ICOT features were observed which for weak coupling looked more like steps in the dI/dV, as expected from second-order theory, whereas for strong coupling nonequilibrium Kondo enhancement was observed (cf. discussion at end of Sec. 4.1). For the experimental data and its theoretical analysis see [62].

In Fig. 10(a) the dI/dV_b (V_g, B) map, calculated from Eq. (64) is plotted where $B = |\mathbf{B}|$ is the strength of the external magnetic field for a fixed gate voltage well within in the Coulomb blockade regime at the experimental temperature. The bright color (left, middle panel) indicates a step in the individual dI/dV_b vs. V_b traces which are illustrated in the right panel (traces taken from middle panel). For a field along the molecule's intrinsic easy z-axis (left panel) a linear Zeeman effect is found in the transport, which is symmetric in B (not shown) with a so-called zero-field splitting (ZFS) remaining at $B = 0$ due to the nonzero anisotropy D-term, see Fig. 9(b). For fields nearly transverse to this axis (angle 71°, middle panel) the Zeeman effect is suppressed and nonlinear. These features, characteristic of a single-molecule magnet with a dominant D-term, were indeed observed in the transport experiments of [62].

We now focus on additional calculations reported in the Supplementary Information for [62], which investigated the possible signatures of spin-tunneling caused by the E-term in Eq. (66) at temperatures lower than the experimental one. We find that the dI/dV_b step sharpens up and shows a slight "overshoot" due to nonequilibrium occupations of the spin states [68] which are correctly included in the calculations using Eq. (64). More interesting is the additional structure which appears in Fig. 10(b) below the ICOT threshold associated with the D-term. These features derive from quantum spin tunneling due to the presence of a sizeable E-term in the model. Their nontrivial magnetic field dependence, shown in Fig. 10(c) for $T = 0.05$ K, ex-

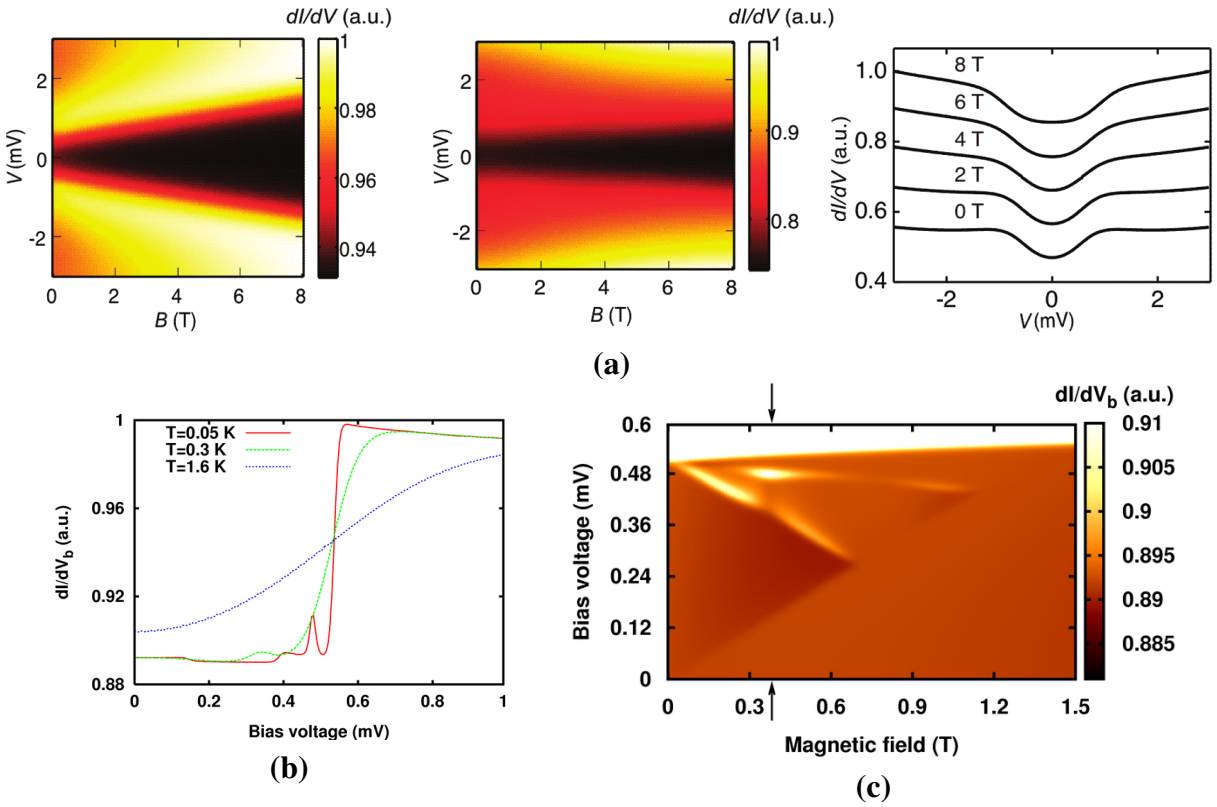


Fig. 10: *Inelastic cotunneling through a single molecule magnet. See description in the text. The arrow in (c) indicates where the red trace in (b) is taken. Figure adapted from [62].*

hibits a Zeeman *splitting* of spin states, indicating that these features are related to states with opposite spin projection, i.e., from the other side of the magnetic anisotropy barrier sketched in Fig. 9(b). Such "subgap" ICOT spectroscopy thus provides direct information about quantum spin tunneling in a single-molecule magnet, and has been observed in STM on magnetic adatoms [69, 70, 71, 72].[16]

In the above examples of second-order perturbation theory we mostly focused on ICOT and COSET spectroscopy of "bare" excitations (electronic, vibrational, or spin related) in the CB regime. However, various other nontrivial effects can occur. For instance, the dependence of molecular quantum-dot energies on the applied voltages may lead to quite unexpected features in the Coulomb blockade regime as observed in [73, 74]. It turns out that these are also captured by the second-order pertubation theory discussed here [75]. Another interesting recent result [37] demonstrates that high-spin quantum dots without any spin-orbit coupling can *acquire* an anisotropy D-term [i.e., given by Eq. (66) with $E = 0$] when coupled to a ferromagnet, and display all the above characteristic ICOT signatures. In such systems, the D-term can thus be generated as a spintronic field by renormalization processes, just like the exchange field (51). A full description of such promising effects, however, requires nonperturbative methods.

Acknowledgments. All collaborators on the cited articles are acknowledged, in particular M. Hell for kindly providing figures and for proofreading the text.

[16] Also here it is important to note that these features do not appear below the ICOT threshold. They only appear below the ICOT threshold *of the model assuming* $E = 0$: the features arise because of the presence of an E-term. This should not be confused with the unphysical appearance of COSET resonances below the ICOT resonances as discussed for the example Fig. 7, see also the related footnote for Fig. 8.

References

[1] M. Leijnse and M. R. Wegewijs, Phys. Rev. B **78**, 235424 (2008).

[2] S. Koller, M. Grifoni, M. Leijnse, and M. R. Wegewijs, Phys. Rev. B **82**, 235307 (2010).

[3] R. B. Saptsov and M. R. Wegewijs, Phys. Rev. B **86**, 235432 (2012).

[4] R. B. Saptsov and M. R. Wegewijs (2013), arXiv:1311.1368.

[5] H. Schoeller, Eur. Phys. Journ. Special Topics **168**, 179 (2009).

[6] J. Eckel, F. Heidrich-Meisner, S. Jakobs, M. Thorwart, M. Pletyukhov, and R. Egger, New. J. Phys. **12**, 043042 (2010).

[7] S. Andergassen, V. Meden, H. Schoeller, J. Splettstoesser, and M. Wegewijs, Nanotechnology **21**, 272001 (2010).

[8] F. Haupt, M. Leijnse, H. L. Calvo, L. Classen, J. Splettstoesser, and M. R. Wegewijs, Phys. Stat. Solidi B **250**, 2315 (2013).

[9] I. O. Kulik and R. I. Shekhter, Sov. Phys. JETP **4**, 308 (1975).

[10] K. K. Likharev, Proc. IEEE **87**, 606 (1999).

[11] J. König, J. Schmid, H. Schoeller, and G. Schön, Phys. Rev. B **54**, 16820 (1996).

[12] U. Fano, Rev. Mod. Phys. **29**, 74 (1957).

[13] C. Timm, Phys. Rev. B **77**, 195416 (2008).

[14] D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures* (Cambridge University Press, 1997), chap. Quantum dots and single electron phenomena, p. 105.

[15] W. G. van der Wiel, S. D. Franceschi, J. M. Elzerman, T. Fujisawa, S. Tarucha, and L. P. Kouwenhoven, Rev. Mod. Phys. **75**, 1 (2003).

[16] C. Timm, Phys. Rev. B **83**, 115416 (2011).

[17] K. Kaasbjerg and K. Flensberg, Nano Lett. **8**, 3809 (2008).

[18] B. Roche, R.-P. Riwar, B. Voisin, E. Dupont-Ferrier, R. Wacquez, M. Vinet, M. Sanquer, J. Splettstoesser, and X. Jehl, Nature Comm. **4**, 1581 (2013).

[19] E. A. Osorio, K. O'Neill, M. R. Wegewijs, N. Stuhr-Hansen, J. Paaske, T. Bjørnholm, and H. S. van der Zant, Nano Lett. **7**, 3336 (2007).

[20] E. Bonet, M. M. Deshmukh, and D. C. Ralph, Phys. Rev. B **65**, 045317 (2002).

[21] H. B. Heersche, Z. de Groot, J. A. Folk, H. S. J. van der Zant, C. Romeike, M. R. Wegewijs, L. Zobbi, D. Barreca, E. Tondello, and A. Cornia, Phys. Rev. Lett. **96**, 206801 (2006).

[22] M.-H. Jo, J. E. Grose, K. Baheti, M. M. Deshmukh, J. J. Sokol, E. M. Rumberger, D. N. Hendrickson, J. R. Long, H. Park, and D. C. Ralph, Nano Lett. **6**, 2014 (2006).

[23] C. T. Chang, J. P. Sethna, A. N. Pasupathy, J. Park, D. Ralph, and P. McEuen, Phys. Rev. B **76**, 045435 (2007), cond-mat/0605671.

[24] J. E. Grose, E. S. Tam, C. Timm, M. Scheloske, B. Ulgut, J. J. Parks, H. D. Abruña, W. Harneit, and D. C. Ralph, Nature Materials **7**, 884 (2008).

[25] A. Eliasen, J. Paaske, K. Flensberg, S. Smerat, M. Leijnse, M. R. Wegewijs, H. I. Jørgensen, M. Monthioux, and J. Nygård, Phys. Rev. B **81**, 155431 (2010).

[26] K. Goß, S. Smerat, M. Leijnse, M. R. Wegewijs, C. M. Schneider, and C. Meyer, Phys. Rev. B **83**, 201403(R) (2011).

[27] K. Goß, M. Leijnse, S. Smerat, M. R. Wegewijs, C. M. Schneider, and C. Meyer, Phys. Rev. B **87**, 035424 (2013).

[28] M. M. E. Baumgärtel, M. Hell, S. Das, and M. R. Wegewijs, Phys. Rev. Lett. **107**, 087202 (2011).

[29] M. Braun, J. König, and J. Martinek, Phys. Rev. B **70**, 195345 (2004).

[30] M. Hell, S. Das, and M. R. Wegewijs, Phys. Rev. B **88**, 115435 (2013).

[31] M. Braun, J. König, and J. Martinek, Superlatt. and Microstruct. **37**, 333 (2005).

[32] M. Braun, J. König, and J. Martinek, EPL **72**, 294 (2005).

[33] M. Braun, J. König, and J. Martinek, Phys. Rev. B **74**(7), 075328 (2006).

[34] M. Braun, J. König, and J. Martinek, *Manipulating Single Spins in Quantum Dots Coupled to Ferromagnetic Leads* (Springer, 2006), Lecture Notes in Physics.

[35] R. Hornberger, S. Koller, G. Begemann, A. Donarini, and M. Grifoni, Phys. Rev. B **77**, 245313 (2008).

[36] B. Sothmann and J. König, Phys. Rev. B **82**, 245319 (2010).

[37] M. Misiorny, M. Hell, and M. R. Wegewijs, Nature Physics **9**, 801 (2013).

[38] B. Wunsch, M. Braun, J. König, and D. Pfannkuche, Phys. Rev. B **72**, 205319 (2005).

[39] A. Donarini, M. Grifoni, and K. Richter, Phys. Rev. Lett. **97**, 166801 (2006).

[40] D. Darau, G. Begemann, A. Donarini, and M. Grifoni, Phys. Rev. B **79**, 235404 (2009).

[41] M. G. Schultz and F. von Oppen, Phys. Rev. B **80**, 033302 (2009).

[42] A. Donarini, G. Begemann, and M. Grifoni, Nano Lett. **9**, 2897 (2009).

[43] A. Donarini, G. Begemann, and M. Grifoni, Phys. Rev. B **82**, 125451 (2010).

[44] A. Donarini, B. Siegert, S. Sobczyk, and M. Grifoni, Phys. Rev. B **86**, 155451 (2012).

[45] S. Sobczyk, A. Donarini, and M. Grifoni, Phys. Rev. B **85**, 205408 (2012).

[46] M. G. Schultz, Phys. Rev. B **82**, 155408 (2010).

[47] S. DeFranceschi, S. Sasaki, J. M. Elzerman, W. G. van der Wiel, S. Tarucha, and L. P. Kouwenhoven, Phys. Rev. Lett. **86**, 878 (2001).

[48] M. Hell, S. Das, B. Sothmann, M. Wegewijs, and J. König, in preparation (2014).

[49] M. Leijnse, M. R. Wegewijs, and M. H. Hettler, Phys. Rev. Lett. **103**, 156803 (2009).

[50] J. Lambe and R. C. Jaklevic, Phys. Rev. **165**, 821 (1968).

[51] R. Schleser, T. Ihn, E. Ruh, K. Ensslin, M. Tews, D. Pfannkuche, D. C. Driscoll, and A. C. Gossard, Phys. Rev. Lett. **94**, 206805 (2005).

[52] J. Paaske, A. Rosch, P. Wölfle, N. Mason, C. M. Marcus, and J. Nygård, Nature Physics **2**, 460 (2006).

[53] H. Schoeller and F. Reininghaus, Phys. Rev. B **80**, 045117 (2009).

[54] E. Sela, H.-S. Sim, Y. Oreg, M. E. Raikh, and F. von Oppen, Phys. Rev. Lett. **100**, 056809 (2008).

[55] D. Boese and H. Schoeller, Eur. Phys. Lett. **54**, 668 (2001).

[56] S. Braig and K. Flensberg, Phys. Rev. B **68**, 205324 (2003).

[57] J. Koch and F. von Oppen, Phys. Rev. Lett. **94**, 206804 (2005).

[58] G. A. Kaat and K. Flensberg, Phys. Rev. B **71**, 155408 (2005).

[59] M. R. Wegewijs and K. C. Nowack, New. J. Phys. **7**, 239 (2005).

[60] J. Koch, F. von Oppen, and A. V. Andreev, Phys. Rev. B **74**, 205438 (2006).

[61] A. K. Hüttel, B. Witkamp, M. Leijnse, M. R. Wegewijs, and H. S. J. van der Zant, Phys. Rev. Lett. **102**, 225501 (2009).

[62] A. S. Zyazin, J. W. van den Berg, E. A. Osorio, H. S. van der Zant, N. P. Konstantinidis, F. May, M. Leijnse, W. Hofstetter, M. R. Wegewijs, C. Danieli, and A. Cornia, Nano Lett. **10**, 3307 (2010).

[63] L. Bogani and W. Wernsdorfer, Nature Materials **7**, 179 (2008).

[64] F. Elste and C. Timm, Phys. Rev. B **73**, 235305 (2006).

[65] M. Misiorny and J. Barnaś, Phys. Rev. B **75**, 134425 (2007).

[66] C. Timm, Phys. Rev. B **76**, 014421 (2007).

[67] M. Misiorny, I. Weymann, and J. Barnaś, Phys. Rev. B **79**, 224420 (2009).

[68] B. Sothmann and J. König, New. J. Phys. **12**, 083028 (2010).

[69] C. F. Hirjibehedin, C. P. Lutz, and A. J. Heinrich, Science **312**, 1021 (2006).

[70] C. F. Hirjibehedin, C.-Y. Lin, A. F. Otte, M. Ternes, C. P. Lutz, B. A. Jones, and A. J. Heinrich, Science **317**, 1199 (2007).

[71] A. Otte, M. Ternes, K. von Bergmann, S. Loth, H. Brune, C. Lutz, C. Hirjibehedin, and A. Heinrich, Nature Physics **4**, 847 (2008).

[72] A. Otte, M. Ternes, S. Loth, C. Lutz, C. Hirjibehedin, and A. Heinrich, Phys. Rev. Lett. **103**, 107203 (2009).

[73] N. Roch, S. Florens, V. Bouchiat, W. Wernsdorfer, and F. Balestro, Nature **453**, 633 (2008).

[74] E. A. Osorio, K. Moth-Poulsen, H. S. J. van der Zant, J. Paaske, P. Hedegård, K. Flensberg, J. Bendix, and T. Bjørnholm, Nano Lett. **10**, 105 (2010).

[75] C. Stevanato, M. Leijnse, K. Flensberg, and J. Paaske, Phys. Rev. B **86**, 165427 (2012).

C 8 Data-driven materials design [1]

Georg K. H. Madsen
ICAMS
Ruhr-Universität Bochum, 44801 Bochum, Germany

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

From the mastering of metallurgy in ancient times to the invention of semiconductors it is clear that the discovery of new materials is a key component in creating wealth and jobs. With the energy challenges facing the world today, new materials for batteries or thermoelectric, photovoltaic and hard-magnetic materials are key components in improving energy efficiency, storage and conversion and thereby reducing our dependence on fossil fuels.
Despite its importance the discovery of new materials is often based on trial and error. Some important recent discoveries illustrates the state of affairs: The compound MgB_2 was known for almost 50 years before its superconductivity was discovered[1] and the discovery of the large thermoelectric (TE) performance of $CsBi_4Te_6$ is best described by the authors: "Recently we reported that $K_2Bi_8S_{13}$ and β-$K_2Bi_8S_{13}$ have promising TE properties (..). When we moved on to investigate corresponding Te analogues with Cs, we obtained an unexpected result: Instead of $Cs_2Bi_8Te_{13}$, we isolated $CsBi_4Te_6$".[2] Putting such discoveries down to pure serendipity would be unfair; It is no coincidence that the same people tend to be "lucky". At the same time it is also clear that there is a strong need for more rational routes to new materials.
In the following I will introduce three different ideas in materials design, namely (i) (ii) , where a model is used to reduce the phase space to be searched, (iii) , where existing data are examined using data-mining techniques to discover new correlations which again can be used to predict new areas of phase-space to be examined. The ideas are introduced by example, and the aim is not to give a comprehensive review of the field, for this the reader is refereed to the literature.[3, 4]

2 Screening thermoelectric materials

While the idea of screening a large number of materials to find certain compositions with desirable properties is conceptually simply, and to a certain degree can be viewed as automated trial and error, it can pose a large technical challenge. It can be achieved experimentally by combinatorial sputtering of materials libraries[5, 6] or computationally in a high-throughput environment.[7, 8] Furthermore, there is an important conceptual challenge in developing descriptors. In this context a descriptor is an easy to determine parameter, such as atomic volume,

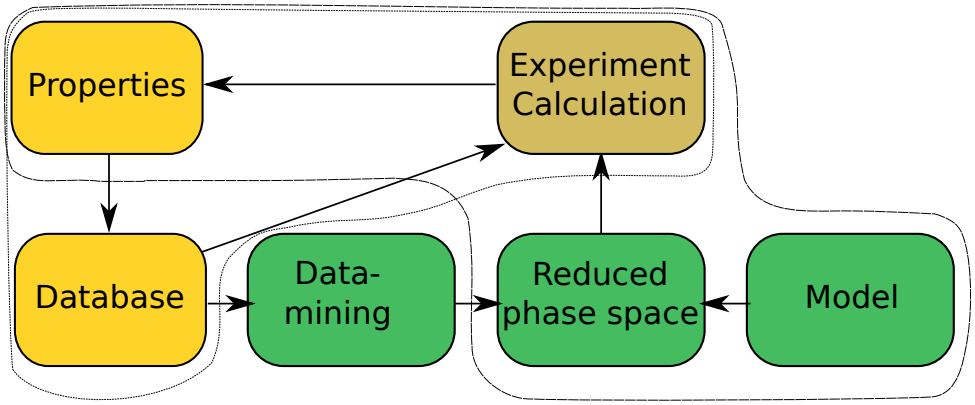


Fig. 1: *Flow diagram of a different high-throughput materials screenings discussed in this paper. The dotted lines correspond to the scheme described in Section 2 and the dashed lines to the scheme described in Section 3*

electronegativities or electron count, which correlates to, or describes, a more complex parameter such as phase stability, band-gap or conductivity. One can view the athermal difference in energy between two phases which can be calculated by density functional theory as a descriptor for phase stability. While it is clear that this, compared to the real phase stability, can suffer from both errors in the applied density functional and from finite temperature effects, there is ample evidence that the DFT formation energies and true experimental stability are closely related.[9, 10]

I will use the high throughput search for new [11] as an early example of the computational high-throughput search for new materials. In the flow diagram of Fig. 1 it corresponds to the dotted area, where a database of known structures is coupled to an automatic calculation band-structures and thermoelectric coefficients. Considering first the descriptor that was optimized in this study, it is clear that it should be based on the dimensionless

$$zT = (\sigma T/\kappa)S^2 \tag{1}$$

which quantified the thermoelectric performance. In Eq. (1), S is the Seebeck coefficient and σ and κ are the electronic and thermal conductivities respectively. The power factor is given by: $PF = S^2\sigma$. S and σ can be written as

$$\sigma = \mathcal{L}^{(0)} \quad , \quad S = \frac{1}{qT}\frac{\mathcal{L}^{(1)}}{\mathcal{L}^{(0)}} \tag{2}$$

where the $\mathcal{L}$ are related to the Onsager coefficients and are formally given by

$$\mathcal{L}^{(\alpha)} = q^2 \int \sigma(\varepsilon)(\varepsilon - \mu)^{\alpha}\left(-\frac{\partial f}{\partial \varepsilon}\right)\mathrm{d}\varepsilon \tag{3}$$

where f is the Fermi-distribution and σ the transport distribution

$$\sigma(\varepsilon) = \sum_n \int v_{n\mathbf{k}} v_{n\mathbf{k}} \tau_{n\mathbf{k}} \delta(\varepsilon - \varepsilon_{n\mathbf{k}}) \frac{\mathbf{dk}}{8\pi^3} \tag{4}$$

The group velocities, $v_{n\mathbf{k}}$ in Eq. (4), can be calculated from the .[12] Furthermore, due to the $\partial f/\partial\varepsilon$ factor only allowing contributions from a narrow energy window, it is reasonable to approximate the relaxation time, $\tau_{n\mathbf{k}}$, with a constant. This approximation means that the relaxation time means that τ cancels in Eq. (2) and we can calculate S on an absolute scale from just the band structure. A further simple consideration can also be made when searching for the optimal zT: κ has both an electronic, κ_e, and an lattice, κ_l, component and the ratio between σT and κ_e is, to a good approximation, given by the Wiedemann-Franz relation: $\kappa_e = L_0\sigma T$. Inserting the Lorentz number, $L_0 = \frac{\pi^2}{3}\left(\frac{k_B}{e}\right)^2 = (156\ \mu\mathrm{V/K})^2$, zT is seen to be limited by the Seebeck coefficient

$$zT = \frac{S^2\sigma T}{\kappa_e + \kappa_l} = \frac{S^2}{L_0 + \kappa_l/(\sigma T)} < \left(\frac{S}{156\mu\mathrm{V/K}}\right)^2 \tag{5}$$

It is thus clear that to obtain a state of the art thermoelectric material with $zT > 1$ one must have a $S > 156\mu$V/K. At the same time $L_0\sigma T$ must be larger than κ_l, which means that the thermal should be small and the electrical conductivity as large as possible. It is impractical to calculate τ and κ_l in a high-throughput fashion. Instead we will rewrite Eq. (5) as

$$zT = \frac{S^2\sigma/\tau}{L_0 T\sigma/\tau + \kappa_l/\tau}T \tag{6}$$

which introduces κ_l/τ as a parameter. Heavily doped semiconductors typically have $\tau = 1 - 2 \cdot 10^{14}$ s. If we use $\kappa_l/\tau = 10^{14}$ W/Kms the resulting zT can be viewed as a descriptor, which obviously does not give the true figure of merit, but gives an absolute scale of a materials thermoelectric potential given that a $\kappa_l = 1 - 2$ W/Km can be reached.
The general validity of this approach has been discussed[13] and it lead to the prediction of as a thermoelectric material[11] that have been experimentally verified.[14]

3 Band structure engineering of new thermoelectric materials

A knowledge based materials design is also a straight forward concept, namely to use our knowledge or understanding about a problem to reduce the phase space to be searched, corresponding to the dashed area in Fig. 1. It is something most of us do intuitively. A simple example would be using the periodic system to find possible dopants. If the aim is to n-dope Si then it is most natural to try the atoms with one extra valence electron, alas the main group V elements, first. In the following I'll try and illustrate how a few basic insights can lead to the design of optimized thermoelectric materials. The discussion is based on the experimental work on $Mg_2Si_xSn_{1-x}$[15] and $PbTe_xSe_{1-x}$[16], where the example below is based on theoretical band structures.
Considering again the transport coefficients, Eq. (3), it is seen that for two channels conducting in parallel the conductivity is simply the sum of the contributions from each channel

$$\mathcal{L}^{(\alpha)} = \mathcal{L}^{(\alpha)'} + \mathcal{L}^{(\alpha)''} \tag{7}$$

So that the Seebeck coefficient is given as

$$S = \frac{1}{qT} \frac{\mathcal{L}^{(1)'} + \mathcal{L}^{(1)''}}{\mathcal{L}^{(0)'} + \mathcal{L}^{(0)''}} \tag{8}$$

Eq. (8) means that if two bands are aligned within the relative narrow energy range defined by the $\partial f/\partial \varepsilon$ factor, Eq. (3), the Seebeck coefficient is weighted sum of the transport coefficients of the two bands. If the two bands are similar the S will be approximately unchanged, The power factor will then be optimized by the conductivity being a sum of the two bands.
Fig. 2 shows the band structure of and . The shape of the bands is very similar, however a close inspection shows that a heavy and light mass band along the $\Gamma - X$ direction have swapped order at the X point. If we make the simple assumption that Si and Sn are chemically similar and this is a pure volume effect then, following Vegards law, one would expect that an alloying of the two compounds would lead an alignment of the two bands and, following Eq. (8), an enhancement of thermoelectric performance. The optimization is illustrated in Fig. 2b.

4 Structure maps

A step beyond the knowledge based materials design would be the use of large amounts of data, e.g. obtained by high-throughput methods, to formulate the rules and correlations automatically. This known as , which can be defined as the process of discovering patterns in data present in substantial quantities. The process must be automatic or semiautomatic and the patterns discovered must be meaningful in that they lead to some advantage.[17]

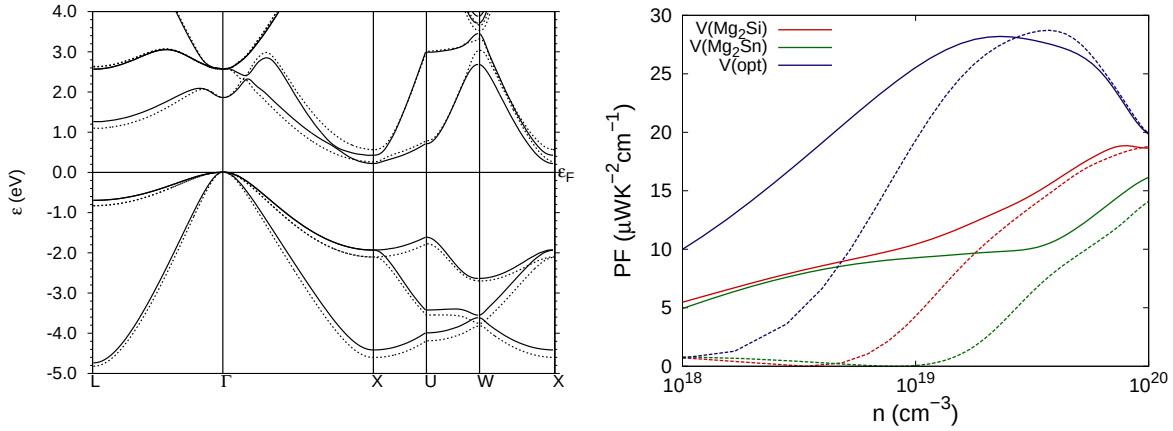


Fig. 2: *a) Band structure of Mg_2Si (full lines) and Mg_2Sn (dotted lines). The conduction bands of Mg_2Sn have been shifted up by 0.39 eV to align them with those of Mg_2Si at the Γ-point. b) Calculated power factor[12] with $\tau = 10^{-14}$ s. The full lines are for Mg_2Si and the dashed lines for Mg_2Sn. The red lines are calculated using the experimental volume of Mg_2Si ($a_0 = 6.36$ Å) and green that of Mg_2Sn ($a_0 = 6.75$ Å). The blue line is at the volume which optimizes the power factor ($a_0 = 6.55$ Å).*

A very early use of this approach is the Pettifor .[18] The Pettifor structure maps order the structures of binary compounds according to a chemical scale, χ, which allows binary alloys A_xB_{1-x} to be mapped onto a two-dimensional coordinate system, with the χ_A on the abscissa and χ_B on the ordinate. It thus correlates the stable structures to a simple one dimensional scale. An interesting point about χ is that it is purely phenomenological and was defined to a posteriori achieve the best structural separation of all the binary A_xB_{1-x} compounds with a given stoichiometry.[19] The effect of ordering structures after χ is shown in Fig. 3 where a small section of the full AB ($x = 1/2$) structure map shown. where it is seen how alloys with similar stable structures tend to cluster together. It illustrates the dual nature of an element like Mg. In some cases it will behave like a dual valent cation like Calcium, its fellow chalcogen, and Eu and form the ionic NaCl structure with Sulphur and in certain cases behave like Zn and not form 1:1 alloys with Si and Ag, whereas Ca and Eu form the CrB structure, or the closely related to the FeB structure.[20] Also shown in Fig. 3 are the Pauling electronegativities and the covalent radii. It is seen that a similar ordering could also be achieved using these two descriptors.

A structure map has some predictive power. If a new alloy is placed on the map, a guess for the unknown stable structure can be obtained by examining the stable structures of the nearby neighbors. This predictive power was quantified by Morgan et al.[21] by introducing a "distance" in chemical scale between a given AB alloy (with chemical scale (χ_A, χ_B)) and other alloys A'B' (with chemical scales $(\chi_{A'}, \chi_{B'})$).

$$d = \sqrt{(\chi_A - \chi_{A'})^2 - (\chi_B - \chi_{B'})^2} \tag{9}$$

The Morgan paper actually uses a integer scale depending on the position in the structure map. If the other alloys are ordered in terms their distance, then there is about 80%-85% probability of finding the correct structure of AB in the top five of the list. If one takes instead just the five overall most common structures they cover only 50 % of the entrances. [21]

5 Outlook

In the present contribution we contrasted the serendipitous discovery of new materials with specific materials designs. Though there is clearly also many serendipitous discoveries in pharmaceuticals, with the story behind being a highly entertaining one, it is still difficult to imagine that a new drug is discovered and first after wards does one start to look for a decease that it can treat. Simple intuition is often a collection of approximate truths and generalizations based on our knowledge. Obviously, it is an advantage if the knowledge can be expressed in a model, which can make quantitative predictions. As these models become more and more precise, more rational materials designs will become possible, which could lead to a new paradigm for the design of new devices can be imagined, where the development of a new material can be thought into the process from the beginning. The examples taken in this contribution has all been based on bulk single crystal materials. It is clear that materials behavior is often influenced by nano- and micro-structure and interfaces within the multi-grain samples and that in the future such effects should be thought into a materials design.

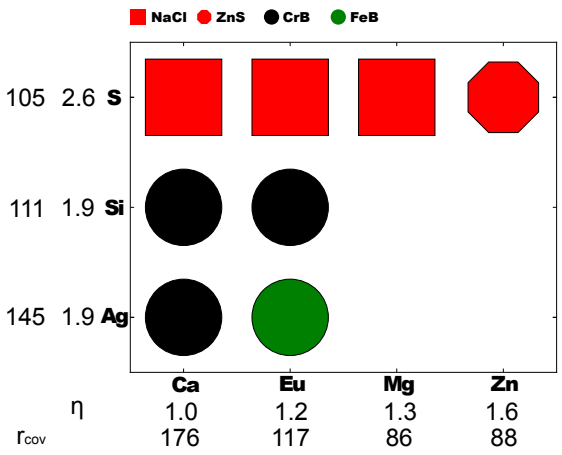


Fig. 3: *Section of full AB structure map.[18] The elements have been ordered according to the chemical scale. Also the Pauling electro-negativity η and the covalent radius in pm of the elements is listed.*

References

[1] J. Nagamatsu, N. Nakagawa, T. Muranaka, Y. Zenitani, and J. Akimitsu, Nature **410**, 63 (2001).

[2] D.-Y. Chung, T. Hogan, P. Brazis, M. Rocci-Lane, C. Kannewurf, M. Bastea, C. Uher, and M. G. Kanatzidis, Science **287**, 1024 (2000).

[3] G. Hautier, A. Jain, and S. Ong, Journal of Materials Science **47**, 7317 (2012).

[4] S. Curtarolo, G. L. W. Hart, M. B. Nardelli, N. Mingo, S. Sanvito, and O. Levy, Nature Materials **12**, 191 (2013).

[5] R. Potyrailo, K. Rajan, K. Stoewe, I. Takeuchi, B. Chisholm, and H. Lam, ACS Combinatorial Science **13**, 579 (2011).

[6] P. J. S. Buenconsejo, A. Siegel, A. Savan, S. Thienhaus, and A. Ludwig, ACS Combinatorial Science **14**, 25 (2012).

[7] A. Jain, G. Hautier, C. J. Moore, S. P. Ong, C. C. Fischer, T. Mueller, K. A. Persson, and G. Ceder, Computational Materials Science **50**, 2295 (2011).

[8] S. Curtarolo, W. Setyawan, S. Wang, J. Xue, K. Yang, R. H. Taylor, L. J. Nelson, G. L. Hart, S. Sanvito, M. Buongiorno-Nardelli, N. Mingo, and O. Levy, Computational Materials Science **58**, 227 (2012).

[9] G. Hautier, S. P. Ong, A. Jain, C. J. Moore, and G. Ceder, Phys. Rev. B **85**, 155208 (2012).

[10] I. Opahle, A. Parma, E. J. McEniry, R. Drautz, and G. K. H. Madsen, New Journal of Physics **15**, 105010 (2013).

[11] G. K. H. Madsen, J. Am. Chem. Soc. **128**, 12140 (2006).

[12] G. K. H. Madsen and D. J. Singh, Comp. Phys. Comm. **175**, 67 (2006).

[13] L. Bjerg, G. K. H. Madsen, and B. B. Iversen, Chem. Mat. **23**, 3907 (2011).

[14] E. S. Toberer, A. F. May, C. J. Scanlon, and G. J. Snyder, J. Appl. Phys. **105**, 063701 (2009).

[15] V. K. Zaitsev, M. I. Fedorov, E. A. Gurieva, I. S. Eremin, P. P. Konstantinov, A. Y. Samunin, and M. V. Vedernikov, Phys. Rev. B **74**, 045207 (2006).

[16] Y. Pei, X. Shi, A. LaLonde, H. Wang, L. Chen, and G. J. Snyder, Nature **473**, 66 (2011).

[17] I. H. Witten and F. Eibe, *Data mining : practical machine learning tools and techniques* (Morgan Kaufmann, San Francisco, U.S.A., 2005).

[18] D. G. Pettifor, *Bonding and Structure of Molecules and Solids* (Clarendon Press, Oxford, UK, 1995).

[19] D. G. Pettifor, J. Phys.: Condens. Matter **15**, V13 (2003).

[20] E. Parth, Acta Crystallographica Section B **32**, 2813 (1976).

[21] D. Morgan, J. Rodgers, and G. Ceder, Journal of Physics: Condensed Matter **15**, 4361 (2003).

D 1 Parallel Computer Architectures [1]

D. Pleiter
Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Since John von Neumann [Neumann(1945)] and others established modern computer processor architecture concepts in the 1940s, the speed at which arithmetic computations could be performed has increased by many orders of magnitude. One ingredient for this increase in performance has been improvements in computer architectures. The other important driver for this development has been the technology for building integrated circuits. Already in the 1960s it had been observed by Gordon Moore that the number of components per integrated circuit increases roughly by a factor 2 per year [Moore(2006)]. A more long-term analysis of the transistor count for Intel microprocessors revealed a doubling every 2 years [Bohr(2007)]. It has become popular to refer to a so-called "Moore's Law", which today is commonly understood as a doubling of the performance every 18 months.

Over some period of time this increase in performance could be achieved by means of increasing the frequency f at which computing devices are clocked. Today all these devices are based on CMOS technology for which dissipated (active) power scales proportional to $V^2 \cdot f$, where V is the voltage. As voltage has to be increased for larger frequencies, power dissipation grows more than linear when f is increased. Since a couple of years frequency scaling has reached its limit. Today's high-performance processors are typically operated in a frequency range $2.5 \lesssim f \lesssim 4\,\text{GHz}$. Further increase of performance can thus largely only be achieved by scaling-out to a larger number of computational devices running in parallel.

High-performance computing today therefore means massively-parallel computing. In this way still further performance increases can be achieved. This is, e.g., documented by the Top500 project,[2] which publishes twice a year an international list of computers which are fastest in terms of floating-point operation throughput while executing the High-Performance LINPACK (HPL) benchmark. In Fig. 1 we show how the (aggregate) performance of the top 1, 10 and 100 systems increased over time. Let us consider the aggregate performance of the top 100 systems in more detail. From June 1993 to November 2013 their performance doubled almost every 14 months on average, now reaching 180 PFlop/s. This corresponds to an performance increase by a factor $2.5 \cdot 10^5$. During that time the number of processor cores increased by roughly a factor 750. This number reflects only a fraction of the increase in parallelism as we observe an increase of parallelism at many levels of the computer architectures, as we will see later.

This lecture is written for computational scientists for which availability of these massively-parallel computing resources are critical for their research. The purpose is to improve the efficient use of such resources by means of better understanding of computer architectures. We will discussed some of the most relevant conceptional aspects which are realised in today's architectures. But we will also have a detailed quantitative look into crucial performance parameters of parallel computer architectures and the underlying technologies.

This lecture starts with a general overview on system hardware architectures and its components in section 2, followed by a discussion of the term performance and an introduction into performance models in section 3. In the following sections we will have a closer look into specific components like the processor (section 4), the memory (section 5) as well as the network (section 7). In section 6 we will introduce a special type of computing devices, so-called accelerators, which play an increasingly important role in high-performance computing. As we will see in section 8, many of the current trends in the development of massively-parallel computers are due to the need for improved power and energy efficiency. In section 9 we will

[2] `http://www.top500.org`

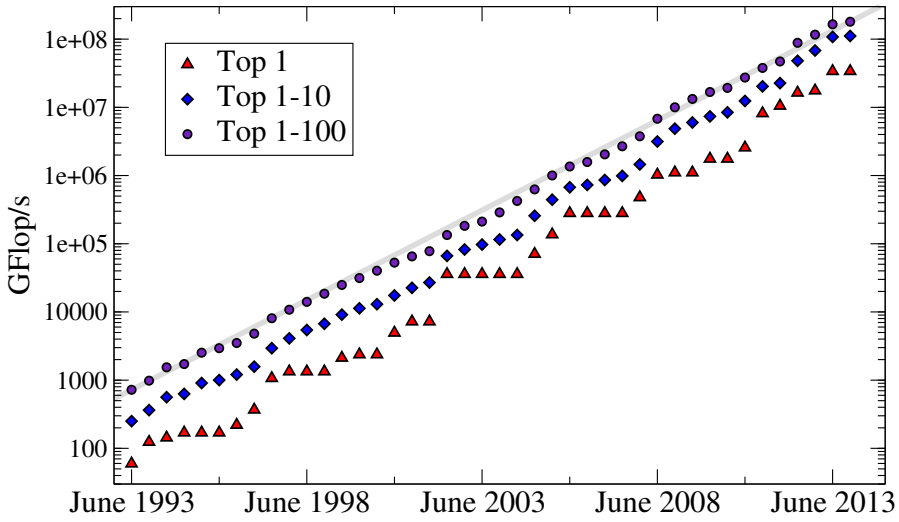


Fig. 1: *The aggregate double-precision floating-point performance achieved by the top 1, 10 and 100 systems on the Top500 list as a function of time. The grey line shows the performance trend for the top 100 systems indicating a doubling of performance every 13.7 months.*

review the trend to increased parallelism, followed in section 10 by a comparison of several high-performance computer architectures, which as of today are intensively used for scientific computations. We will close with an outlook in section 11.

2 Abstracted View on System Hardware Architectures

Today's processor architectures are surprisingly similar to the architecture proposed by von Neumann [Neumann(1945)]. We thus use the latter, significantly simpler architecture as a model to introduce concepts which are helpful to understand modern architectures. In Fig. 2 we show a variant of the Von Neumann architecture. The memory unit *MEM* holds both data and instructions which are processed by the central processing unit *C*. The latter consists of a central control unit *CC* which directs operations, i.e. fetching of instructions and operands, dispatch of instructions and storing results. The central arithmetic unit *CA* executes the arithmetic instructions. Communication with the processor requires the availability of input and output units. All these units are interconnected by a central bus.

Let us assume that data and program have been loaded into the processor's memory and execution of instructions can start. This means that the first instruction can be fetched followed by a fetch of the operands. To hold instruction and operands in the central processing unit small pieces of (fast) memory, called *registers*, are required. Once instruction and input operations are available for execution, the instruction can be issued and results written back to memory. A program counter, which points to the current instruction, then needs to be incremented before the next instruction is loaded. In a very simple program this cycle continues until all instructions have been processed.

This architecture has a crucial limitation, called the Von Neumann bottleneck. Both instructions and data are kept in the same memory and have to be moved along the same data path to the central processing unit resulting in potential congestion on the internal bus. This problem can

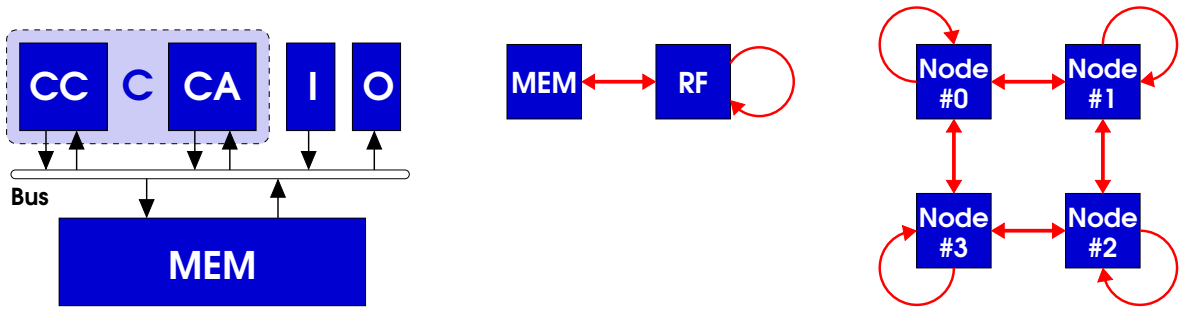


Fig. 2: *On the left a version of the Von Neumann architecture is shown. The central processing unit* C *consists of the central control and arithmetic units* CC *and* CA*, respectively. A common bus connects* C *to the memory* MEM *as well as to the input and output units (*I *and* O*). The middle figure shows an abstract processor architecture model comprising 2 storage devices: The memory* MEM *and a register file* RF*. The graph on the right represents a parallel architecture comprising 4 nodes connected by network links in a ring topology.*

be mitigated by modifying the architecture such that instructions are kept in a separate memory unit directly connected to the central processing unit.[3] Now instructions and operands can in principle be loaded in parallel.
This short digression on modified Von Neumann architectures highlights the importance of data transport within a system hardware architecture. In the following we introduce an abstract system hardware architecture model which focusses on this aspect. The model comprises two types of devices:

Storage devices: Devices which can store data and are characterized by their capacity.

Transport or processing devices: Devices which read data from one storage device, process the data and write the result to another or the same storage device. The data does not have to be processed but may just be copied from one device to another. The performance of the device is in particular characterized by throughput or bandwidth, i.e. the speed of data processing or transport, respectively.

The model can naturally be represented by a graph where the vertices and edges refer to storage and transport devices, respectively. Fig. 2 (middle) shows such a graph for a very simple processor architecture comprising of 2 storage devices: the main memory and a set of registers, called *register file*. These are connected by a memory bus through which data can be read from memory to the register file or written back to memory. The register file is connected to a processing pipeline which takes its input operands from the register file and writes results back to the same register file. The processing pipeline may, e.g., be capable to perform fused multiply-add operations.
It should be noted that the abstraction level of this model can be adjusted. The previous example describes a processor architecture and is relatively close to the actual hardware architecture of very simple processors. Let us consider a parallel architecture with multiple nodes where each node has potentially a complex internal structure, e.g., comprising multiple processors with separate memory units attached to each of them. Each of the nodes could be modeled as a single storage device with one or more processing devices attached, which read data from the

[3] This modified architecture is often called Harvard architecture.

storage device, processes it and write it back to the storage device. The different storage devices, i.e. nodes, are connected by transport devices representing network links. Fig. 2 (right) shows a graphical representation of such a machine with 4 nodes interconnected in a ring topology.

3 Performance and Performance Models

Before we continue on computer architectures with the goal of obtaining a better understanding of the performance a computational scientist might expect for a given application we first have a closer look on different aspects of the term performance.
Here we assume that the user of a high-performance computing system aims on minimizing *time-to-solution* Δt for a given *problem size*, also called *workload*. Time-to-solution refers to the time needed to execute an application, ignoring the efforts required to implement, port or optimize an application, which depending on the complexity of the application may become relevant, too.[4] In future also other performance metrics like energy-to-solution for a given workload might become relevant.
Start and end of application execution can be seen as events, i.e. points of time where something happens. Let us refer to these events by t_{start} and t_{end} and write time-to-solution as $\Delta t = t_{\mathrm{end}} - t_{\mathrm{start}}$. The problem size refers to all parameters which affect the efforts required to solve a computational problem. For Density Function Theory (DFT) calculations this would, e.g., include both the number of nuclei as well as their position.
Any application can be decomposed into a set of computational tasks that need to be solved. An example for such a task, which can be found in many computational science problems, is solving a linear set of equations, e.g. $A\,x = b$. Within a single application such a task may have to be executed repeatedly for different matrices A and different input vectors b. Depending on the properties of this linear equation different numerical algorithms may be applied. Let us assume iterative methods like the Conjugate Gradient method to be a good choice. In this case the numerical task of solving this equation can be further decomposed in smaller subtasks like matrix-vector multiplications, scalar products and linear combinations. While application developers may organise the implementation of tasks (or subtasks) in separate functions or subroutines, this is not relevant for the theoretical concept formulated here.
In practice, it often can be observed that a large fraction of the total time-to-solution is spent in just a few tasks or subtasks. A set of such tasks, which are executed in sequence, form a *performance critical region*, also called computational kernel.
It is important to distinguish the following types of performance:

Machine performance: This metric refers to the amount of work that our computer architecture performs within a given time unit on a given set of resources (e.g. number of nodes). The work here is defined in terms of hardware related units. An example is the number floating-point operations which the machine is able to execute per second while running a given computational task.

Algorithmic performance: The algorithmic performance quantifies the number of steps per computational task in algorithmic terms.In case of the Conjugate Gradient algorithm or other Krylov-space based iterative methods the amount work can for instance be measured in terms of number of iterations required to obtain a solution within a given precision.

[4] See [Wienke et al.(2013)] for an attempt to take costs of application development into account.

The performance of an application for a given problem size depends on both, the machine as well as the algorithmic performance. The application developer is faced with the challenge to decide on the most promising strategy to increase the performance of an application, i.e. reduce time-to-solution, as machine performance may deteriorate when using a different algorithm with better algorithmic performance and vice versa. The better choice may also be problem size dependent.

While we focus here on measuring machine performance in terms of time-to-solution, it should be noted that for more detailed performance analysis not only the number of clock cycles spent in a certain task or subtask but also other counters, which, e.g., count the number of floating-point operations or load/store operations, can provide important insight. Modern processors comprise performance monitoring units that enable various hardware events to be counted (see, e.g., [Terpstra et al.(2010)] for more details).

As implementing algorithms just to measure time-to-solution is often not affordable, performance models can become important as they allow to make predictions of machine performance. Performance models also allow to understand how the observed performance at system level relates to performance at device level and how, e.g., the performance of a transport device like the network impacts overall performance of an application.

In the previous section we introduced a simple model where we represented a computer architecture by a graph where the nodes and vertices are storage devices and transport or processing devices, respectively. Each task k which is executed by this computer architecture implies that information is transferred from one storage device x to another storage device y. The amount of information is given by the *information exchange* $I^k_{x,y}(W)$, which is a function of the problem size W [Bilardi et al.(2005)].

Let us consider a practical example on how to determine the information exchange where an array a is copied on an array b, i.e. $b_i \leftarrow a_i$. Both arrays comprise N double-precision floating-point numbers of size 64 bit. When executing this task, N elements of a have to be loaded from memory to register file and subsequently N elements of b have to be transferred back to memory. The problem size can thus be parametrized by the input parameter N. In this example we find

$$I^{\mathrm{cp}}_{\mathrm{mem,rf}}(N) = I^{\mathrm{cp}}_{\mathrm{rf,mem}}(N) = 8N\,\mathrm{Byte}. \tag{1}$$

When loading N array elements from memory a certain amount of time will elapse between the event "load instruction issued" until the first data arrives. This time is called *start-up latency*. Once the first element of array a arrives at the register file, the next data will follow. This behaviour can in first approximation be parametrized by the following linear ansatz:

$$\Delta t_{\mathrm{cp}}(N) = 2\left(\lambda_{\mathrm{mem}} + \frac{I^{\mathrm{cp}}(N)}{\beta_{\mathrm{mem}}}\right), \tag{2}$$

where λ_{mem} and β_{mem} are the start-up latency and *asymptotic bandwidth*, respectively. The factor 2 accounts for data being read and written.[5] The asymptotic bandwidth should be distinguished from the effective bandwidth which in our case is

$$b_{\mathrm{mem}}(N) = \frac{2\,I^{\mathrm{cp}}(N)}{\Delta t_{\mathrm{cp}}} = \left(\frac{\lambda_{\mathrm{mem}}}{I^{\mathrm{cp}}(N)} + \frac{1}{\beta_{\mathrm{mem}}}\right)^{-1}, \tag{3}$$

[5]Here we assume start-up latency and bandwidth for memory read and write operations to be the same. This is not alway a valid assumption.

i.e. $b_{\mathrm{mem}}(N) \leq \beta_{\mathrm{mem}}$. The effective bandwidth depends, unlike the asymptotic bandwidth, on the problem size (here N). For small problem sizes the effective bandwidth strongly depends on the start-up latency, while for large N the effective bandwidth approaches the asymptotic bandwidth:

$$\lim_{N\to\infty} b_{\mathrm{mem}}(N) = \beta_{\mathrm{mem}}. \tag{4}$$

The asymptotic bandwidth may be put in relation with the *nominal bandwidth* of a given transport device, here the nominal memory bandwidth B_{mem}. In the simple processor architecture considered here (with, e.g., caches being absent) we have $\beta_{\mathrm{mem}} \leq B_{\mathrm{mem}}$.

Despite the fact that the underlying hardware mechanisms in modern processors are very complex (as we will see later), this simple *latency-bandwidth model* describes the measured timings often well (or at least sufficiently well).[6] Once start-up latency and asymptotic bandwidth have been determined for a given hardware architecture, then the knowledge of the information exchange function, which depends on the computational task of interest and the problem size, allows for a prediction of the execution time based on these simple hardware and performance models.

As a slightly more complex example let us consider the scalar product $c \leftarrow \sum_i a_i\, b_i$. Here N elements of a and b have to be loaded from memory, multiplied with each other and added to the scalar variable c. The latter variable is assumed to be kept in a register, i.e. it does not have to be transferred before writing the final result, an operation which we will ignore as we assume $N \gg 1$. We thus have two information exchange functions describing the memory-to-register file transfer and the processing of the data in the pipeline for fused multiply-add operations:

$$I^{\mathrm{sp}}_{\mathrm{mem,rf}}(N) = 16N\,\mathrm{Byte}, \tag{5}$$

$$I^{\mathrm{sp}}_{\mathrm{rf,rf}}(N) = 2N\,\mathrm{Flop}. \tag{6}$$

Like for the copy operation, the information exchange functions may be used to predict the time needed for loading and processing data, Δt_{mem} and Δt_{proc}, based on the latency-bandwidth model. Making the (realistic) assumption that loading and processing of data can be overlapped, the total execution time for this task can be estimated as $\Delta t_{\mathrm{sp}} = \max(\Delta t_{\mathrm{mem}}, \Delta t_{\mathrm{proc}})$.

For different system hardware architectures Δt_{mem} and Δt_{proc} may differ significantly, depending on how the floating-point processing pipeline's throughput B_{fp} relates to the available memory bandwidth B_{mem}. The ratio of required number of arithmetic operations versus the amount of data, which needs to be transferred, is called *arithmetic intensity* [Harris(2005)]. For the scalar product we can express the arithmetic intensity in terms of the already determined information exchange functions:

$$\mathrm{AI}_{\mathrm{sp}} = \frac{I^{\mathrm{sp}}_{\mathrm{rf,rf}}(N)}{I^{\mathrm{sp}}_{\mathrm{mem,rf}}(N)} = \frac{1}{8}\frac{\mathrm{Flop}}{\mathrm{Byte}}. \tag{7}$$

For the system hardware architecture being balanced, i.e. both memory interface and processing pipeline utilization being approximately equal, we would like to have[7]

$$\frac{B_{\mathrm{fp}}}{B_{\mathrm{mem}}} \simeq \frac{1}{8}\frac{\mathrm{Flop/s}}{\mathrm{Byte/s}}. \tag{8}$$

[6] In case of more complex network protocols, where depending on the amount of data to be communicated different protocols are used, more sophisticated models may be required, like the LogP model [Culler et al.(1993)].

[7] Using the asymptotic bandwidth or throughput (here: β_{mem} and β_{fp}) may be a better choice than the nominal bandwidth or throughput (here: B_{mem} and B_{fp}). While nominal bandwidth figures can be extracted from datasheets, estimates of asymptotic bandwidth are typically more difficult to obtain and may require, e.g., execution of micro-benchmarks.

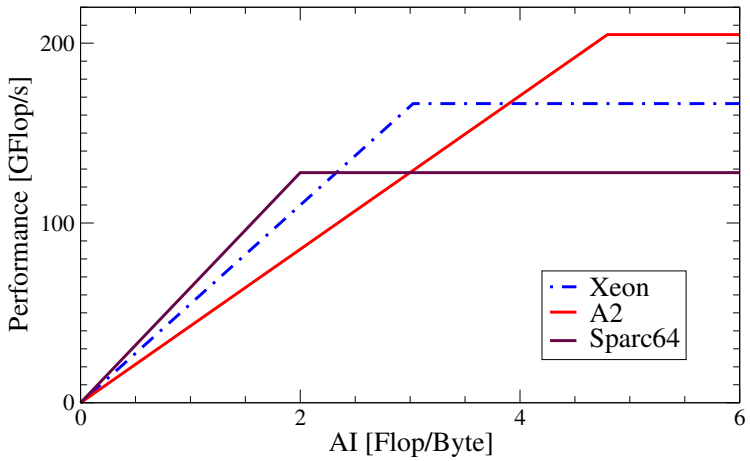


Fig. 3: *Roofline model for the processors listed in Table 1. For small AI (here: AI$\lesssim$ 2) the maximum attainable performance is limited by memory bandwidth, while for large AI (here: AI$\gtrsim$ 4) the performance is limited by floating-point operation throughput. It should be noted that for this (worst case) analysis the existence of caches has been ignored.*

A different look on arithmetic intensity is provided by the roofline model [Williams et al.(2009)]. For small arithmetic intensity AI the floating-point performance is limited by the memory bandwidth. The performance limit increases with larger AI until the point is reached where the throughput of the processing pipeline(s) becomes the limiting factor. In Fig. 3 we show the roofline for different modern server processors. As can easily be seen, a scalar product of two arrays of double-precision floating-point numbers with $\mathrm{AI} = 1/8$ would be strongly bandwidth limited on these processor architectures. Relative to the peak floating-point performance processing capabilities we can not expect to see a sustained performance of more than 6.3 %, even on the for this task presumably most suitable processor, the Sparc64 VIIIfx.

4 Processor Core Architectures

When introducing the Von Neumann architecture we encountered a single central processing unit. Modern processors comprise multiple such units, which we in the following call *processor cores*. Here the instructions of an application are executed. Over the last decades significant advances have been made optimizing instruction throughput, i.e. the number of instructions per cycle (IPC). While logically instructions are executed sequentially, in the order as they appear in the program, at hardware level they may be executed in parallel and/or out-of-order.
All processors, whose hardware parameters are listed in Table 1, comprise cores with multiple execution pipelines and are capable of processing multiple instructions in parallel. Processor cores which implement such instruction-level parallelism (ILP) , and thus allow for $\mathrm{IPC} > 1$, are called superscalar. The number of execution units can vary significantly between different processor core architectures as can be seen from Table 1. The Sparc64 VIIIfx core architecture [Maruyama et al.(2010)] comprises a relatively large number of 8 processing pipelines: 2 for integer instructions, 2 for load/store instructions and 4 for floating-point instructions. The hardware, which decides when which instruction is executed, has to ensure the correctness of instruction execution.
With multiple pipelines being available, it is possible that the order in which instructions are

	Xeon E5-2650	Blue Gene/Q A2	Sparc64 VIIIfx
Core clock speed [GHz]	2.6	1.6	2
Floating-point peak B_{fp} [GFlop/s]	166.4	204.8	128
ISA	x86-64	Power	Sparc-V9
SIMD ISA/width (bits)	AVX/256	QPX/256	HPC-ACE/128
Number of exec. pipelines	5	2	8
Instruction processing	out-of-order	in-order	out-of-order
SMT	2-way	4-way	–
N_{cores}	8	16+1	8
L1 data cache per core [kiByte]	32	16	32
L1 instr. cache per core [kiByte]	32	16	32
L2 cache [MiByte]	2	32	5
L3 cache [MiByte]	20	–	–
Memory bandwidth B_{mem} [GByte/s]	51.2	42.7	64
Memory capacity C_{mem} [GiByte]	O(10-100)	16	16

Table 1: *Hardware parameters of processors used in today's HPC systems: Xeon E5-2650 (Sandybridge) [Intel(2013)] from Intel, Blue Gene/Q A2 [Aho et al.(2013)] from IBM, Sparc64 VIIIfx [Maruyama et al.(2010)] from Fujitsu. The core clock speed refers to the default clock speed. Throughout this lecture we adopt the convention that* $1\,\text{kByte} = 10^3\,\text{Byte}$ *and* $1\,\text{kiByte} = 2^{10}\,\text{Byte} = 1024\,\text{Byte}$. *All 3 processors are programmed using different instruction set architectures (ISA).*

executed differs with respect to program order. This can happen if the execution of instruction i has to be stalled, e.g. because of missing input data, while a later instruction j is ready for execution in a different pipeline. In this way instruction throughput, i.e. average IPC, increases. However, if instructions i and j are to be executed by the same execution unit then this change of order is not possible, at least for processor core architectures where instructions are strictly processed in order. Many processor core architectures meanwhile support out-of-order execution of instructions with different maximum distance at which two instructions can be re-ordered.
Another technique for improving the utilization of processing pipelines is Simultaneous Multi-Threading (SMT). Instead of executing only the instructions of a single software thread, processor core architectures may support concurrent execution of multiple software threads. As the probability increases for at least one of the threads providing instructions ready for execution, the average IPC is likely to increase. Table 1 compares SMT support for different processors. The A2 core used in Blue Gene/Q supports a particular large number of hardware threads to compensate for the lacking support of out-of-order execution of instructions. While SMT may help to increase IPC, it may also lead to increased pressure on resources shared by several threads, e.g. register file.[8] Higher utilization of such potentially scarce resources can have negative impact on performance.
In all of these processors an increasingly often used strategy for enhancing operation throughput has been applied: SIMD (Single Instruction Multiple Data) instructions apply the same operations to multiple data. Let us, e.g., consider a double-precision floating-point fused multiply-add

[8] In the A2 core each thread does have a separate register file.

instruction which implements the operation $d \leftarrow a + b \cdot c$, where a, b, c and d are 64-bit floating-point numbers. A 4-way SIMD version of this instruction would take 3 256-bit input operands and generate a single 256-bit output:

$$\begin{pmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \end{pmatrix} \leftarrow \begin{pmatrix} a_0 + b_0 \cdot c_0 \\ a_1 + b_1 \cdot c_1 \\ a_2 + b_2 \cdot c_2 \\ a_3 + b_3 \cdot c_3 \end{pmatrix} \tag{9}$$

While these SIMD instructions enable a significant increase of the operation throughput at moderate hardware costs, the utilization of such hardware capabilities may be challenging. Although the ability of compilers to perform automatic vectorization is improving, often the user has to take care of explicit SIMD programming[9] or at least provide a suitable data layout. SIMD load operations can improve memory access as more data is loaded using a single instruction. On some architectures, however, address alignment requirements, the possible need of gathering data from different memory locations into a single vector and other complicacies may limit the usability of SIMD instructions.

5 Memory Architectures

Within a system hardware architecture the main performance parameters of integrated memory devices are the following: capacity C_{mem}, bandwidth B_{mem} and access latency λ_{mem}. The performance of the currently most widely used memory technology, DDR-SDRAM (Double Data Rate Synchronous Dynamic Random Access Memory), and similar technologies has mainly increased in terms of capacity. Improvements in terms of bandwidth and latency have been very moderate. If external SDRAM based memory devices would be the only memory available to the processor then the performance of almost any scientific application would be severely limited by memory bandwidth and access latencies.
An important empirical observation helps improving processor hardware architectures: Applications tend to reuse data (and instructions) they have used recently. Therefore data localities exist which can be exploited. These localities can be classified in the following way:

Temporal locality: Recently accessed items are likely to be accessed again in the near future.

Spatial locality: Items stored at nearby addresses in memory tend to be referenced close together in time.

As an example let us consider a very simple, 1-dimensional stencil computation:

$$b_i \leftarrow c_- \cdot a_{i-1} + c_+ \cdot a_{i+1}. \tag{10}$$

To compute b_i we have to read 2 elements of array a which are stored at memory addresses which are very close in memory (in this example the distance is 2 times the size of an element of a). When computing b_{i+2} then a_{i+1} is re-used, i.e. the example also exposes temporal locality. Data locality can be exploited to enhance performance by inserting a storage device between register file and main memory which provides high bandwidth and low latency towards the

[9]To support SIMD instruction set architectures, like SSE, AVX, QPX, HPC-ACE, many compilers provide suitable extensions.

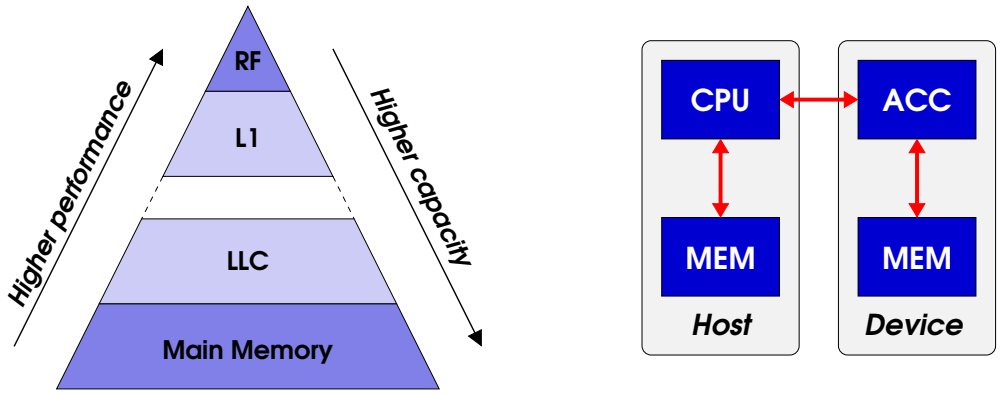


Fig. 4: *A schematic view on a typical processor memory hierarchy is shown in the left figure. On the right a schematic view on an accelerator architecture is shown comprising a host with a single processor as well as a device with a single compute accelerator.*

processor core. High temporal locality means that it becomes beneficial holding a copy of the data in this faster memory device as it is likely that this data is going to be re-used. Transport of data in larger chunks over the memory bus helps to improve bus utilization. High spatial locality increases the probability of all data loaded into the processor to be used, i.e. not only the initially requested data.

This fast memory may either be managed by software or by hardware. The latter is much more common. All processors used in today's HPC architectures comprise even multiple hardware managed caches, i.e. data is moved from memory through 2-4 caches before it arrives at the processing pipelines. The different caches (or more generally speaking: storage devices) are typically arranged in a hierarchical way and thus the term *memory hierarchy* has been established.[10] Such a hierarchy is shown in Fig. 4 (left).

When a load instruction is executed then first the highest cache level, by convention called L1 cache, is checked on whether the requested data is available, i.e. either a cache hit or miss occurs. In the latter case an attempt is made to fetch the data from the next cache level. Only if the *last-level cache* (LLC) does not contain a valid copy of the requested data then data is read from external memory.

The highest cache level provides highest performance, i.e. highest bandwidth and smallest latency, but also the smallest capacity, while the main memory at the bottom of this hierarchy is relatively slow but provides the largest capacity. To make this statement more quantitative: For typical systems used for high-performance computing systems as of today the memory capacity increases from $O(10)$ kiByte at the L1 cache level to $O(10-100)$ GiByte at the main memory level. While it takes $O(1)$ ns to fetch data from the L1 cache, the latency to the main memory directly attached to the processor is $O(100)$ ns. With processor cores running at a core clock speed $2.5 \lesssim f \lesssim 4$ GHz this means that it in the worst case it may take $O(1000)$ clock cycles from the point of time when a load instruction is executed until the data arrives at the processing pipelines.

At the highest, L1 cache level there are typically separate caches for data and instructions to mitigate the effects of the Von Neumann bottleneck, while at lower levels caches are used both

[10] It is, however, important to note that more complex memory architectures play a role in high-performance computing. See, e.g., architectures comprising of accelerator devices (see Fig. 4, right).

for data and instructions. Caches, which are used by a single processor core only, are called private. In processors comprising multiple cores higher cache levels, in particular the LLC are shared. Selected parameters describing the memory hierarchy for server processors used in current HPC systems are shown in Table 1.
Cache misses can be categorized depending on the circumstances under which the cache miss occurs. *Compulsory cache misses* occur when data has to be loaded for the very first time to the cache. The data stays there until it is evicted. This may happen either because the cache became full or because data from another memory address is loaded, which is mapped to the same location in cache. In these cases consecutive *capacity cache misses* or *conflict cache misses* may occur, respectively.
How efficiently the cache hierarchy is used depends on the numerical task, its implementation as well as the problem size. For small problem sizes data reuse might be easier to realise as for larger problem sizes when temporal distance until data is reused becomes too large, i.e. the probability of a data item being evicted from cache before it could be reused becomes large. For a given numerical task and problem size the number of compulsory cache misses cannot be reduced. Changes in implementation and data layout can, however, have significant impact on capacity and conflict cache misses. Strategies on how to optimize for cache utilization will be a topic of later lectures (also see [Kowarschik and Weiß(2003), Hager and Wellein(2010)]).

6 Compute Accelerators

Processor core architectures as presented in an earlier section have become not only very powerful but also complex and costly, e.g. in terms of integrated circuit resources or power consumption. This limits the number of processing units which can be integrated within a single die. For certain compute intensive tasks it may, however, be sufficient to provide very simple compute units. By integrating a very large number of such compute units within a single integrated circuit very large compute performance is achievable. Such devices are often called compute accelerators as they are tightly coupled with a processor and cannot be operated stand-alone. This architectural feature may change in the future.
One may view an architecture comprising processors and compute accelerators as heterogeneous architectures comprising a smaller set of complex, heavy-weight compute cores and a large number of simple, light-weight compute units. While this concept is likely to play an even more important role in high-performance computing in the future, the way how these different compute units are integrated may change quite significantly.
In such heterogeneous architectures the memory hierarchy is (and will be) typically more complex than discussed so far. In Fig. 4 (right) we show a typical node architecture comprising of a host part which corresponds to the architectures discussed earlier: a processor with attached external memory and typically several levels of caches, which have not been shown here. Additionally there is a similar looking device part comprising an accelerator device with attached external memory. Accelerator devices do also comprise multiple levels of cache. Data which needs to be moved from host to device memory (or vice versa) has to be transferred over 2 memory buses as well as a link connecting processor and accelerator. Today this link is typically based on a technology called PCI Express (PCIe) with a link bandwidth $B_{\mathrm{link}} \simeq 8-16$ GByte/s. If we compare this number to the memory bandwidth numbers listed in Tables 1 and 2 then we observe B_{link} to be about 10-50 times smaller than B_{mem}, i.e. special care needs to be taken in order to minimize data transfer over this link.

	K20	K40	Xeon Phi 7120
Vendor	NVIDIA	NVIDIA	Intel
Architecture codename	Kepler	Kepler	Knights Corner
Core clock speed [GHz]	0.71	0.74	1.24
Number of cores/SMs	13	15	61
DP floating-point peak $B_{\mathrm{fp,DP}}$ [TFlop/s]	1.2	1.4	1.2
SP floating-point peak $B_{\mathrm{fp,SP}}$ [TFlop/s]	3.5	4.3	2.4
Memory bandwidth B_{mem} [GByte/s]	208	288	352
PCIe link bandwidth B_{link} [GByte/s]	8	15.7	8

Table 2: *Hardware parameters of selected high-end compute accelerators.*

Today the most widely used type of compute accelerators are GPUs. More details on GPU architectures and on how to program them will be presented in a later lecture. The architecture of these accelerators, which had been originally designed and are still mainly used for graphics processing, differs significantly from processor architectures discussed so far. The entities which may be considered to be the pendants of processor cores are called *streaming multiprocessors (SM)* which are capable of performing significantly more operations in parallel compared to processor cores. Each SM of the GPUs listed in Table 2 can perform 384 single-precision floating-point operations per clock cycle. By means of this huge amount of parallelism these devices can provide an order of magnitude more compute performance compared to processors, although they operate at a significantly lower clock speed. Whether this high compute capability can be efficiently exploited depends, however, strongly on the kind of computational task. For instance, it should be implementable in such a way that a very large number of threads can execute a significant number of instructions independently of each other.
A different strategy is pursued in Intel's Many Core Architecture (MIC) which has been implemented in Intel's Xeon Phi devices. Here a large number of about 60 standard, less powerful processor cores is used. Each of the cores is able to execute very wide SIMD instructions such that each core of the Xeon Phi device shown in Table 2 is able to perform 32 single-precision or 16 double-precision floating-point operations per clock cycle.
Table 2 shows a set of hardware performance numbers for some selected accelerator devices. While the GPU and MIC architectures differ significantly, they do have some common features as well as features which distinguish them from standard processors, e.g. those listed in Table 1. The floating-point performance as well as the memory bandwidth is about 5-10 times larger. All devices run at relatively low clock speed, but are able to execute thousands of floating-point operations per clock cycle. The bandwidth of the link connecting the accelerator device with the host processor is similar and small compared to the memory bandwidth.

7 Network Architectures

As we have seen in the previous section, using compute accelerators a single compute node can provide floating-point processing capabilities at the TFlop/s scale. For high-end computing systems with compute capabilities in the PFlop/s range thousands of nodes have to be integrated in a fast network.

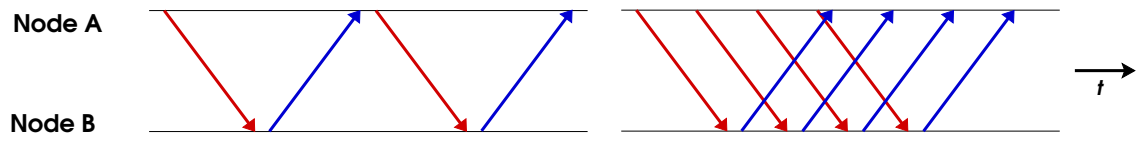


Fig. 5: *The left figure shows a schematic time diagram for a single series of ping-pong communications. In the right figure multiple such series are overlapped to hide communication latency and improve link utilization.*

All state-of-the-art networks are based on point-to-point links which connect compute nodes and/or network devices like switches. Messages which are sent from one node thus usually have to traverse multiple links to reach the receiving node. The most important parameters which describe the performance of such links are bandwidth and latency. While bandwidth continues to improve, further reduction of latency is much more difficult to achieve. Today typical bandwidth performance is $B_{\rm link} = O(1)$ GByte/s, while latency stagnates at $\lambda_{\rm link} = O(100)$ ns.The end-to-end latency of communication at user level is usually $> 1\,\mu$s (frequently even $\gtrsim 10\,\mu$s) because of time needed to communicate between processor and network device plus software overheads.
In terms of core clock cycles the time required to complete a single data transport from one node to the other thus becomes huge. For computational tasks involving significant amounts of small data transfers, i.e. transfers for which execution time is dominated by start-up latency, application developers have to take care of this network architecture limitation. A common strategy, which can be applied for many cases involving fine-grained data transfer, is latency hiding through pipelining. Let us consider a simple pattern of communication between 2 nodes A and B called ping-pong. Node A starts sending a message to B, once this message has been received B sends a message to A, and so on. A schematic version of the time diagram is shown in Fig. 5 (left). Following a pipelining principle the network link connecting the nodes can be used in a much more efficient way if multiple ping-pong communication patterns can be overlapped. As indicated in Fig. 5 (right) a larger number of communications can be performed within the same amount of time in this case. Instead of overlapping one communication operation with other communication operations one can also try to overlap the time between sending and receiving data by performing other work, e.g. computations. The described strategy is commonly called *latency hiding*.
Another key parameter is the network topology. For different communication patterns a given network topology may be more or less suitable. In this lecture we limit us to the following aspects:

Network diameter: The diameter is the maximum distance between 2 nodes. The larger the diameter the more hops may be required to perform a point-to-point communication between nodes and the start-up latency for each communication will thus increase. This also affects the performance of collective operations as a small amount of data has to be collected from all nodes (e.g., in case of a global sum) and/or distributed to all nodes (e.g., in case of a broadcast operation).

Bi-section bandwidth: The bi-section bandwidth refers to the aggregate bandwidth of all links connecting two equal halves of the machine. This metric is important for all-to-all com-

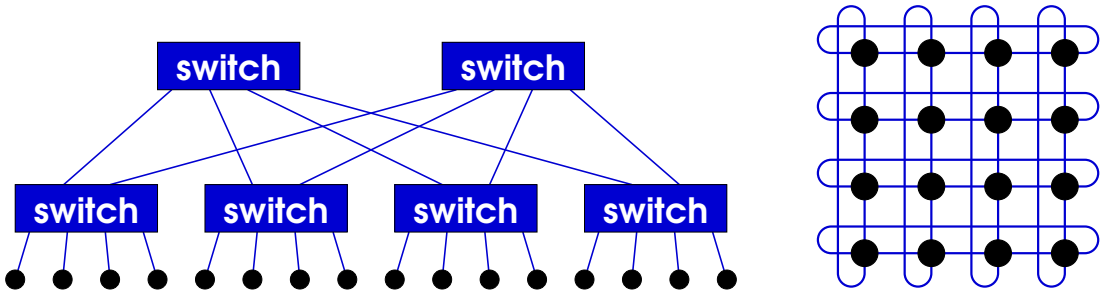


Fig. 6: *Commonly used network topologies: fat-tree (left) and torus (right). The filled circles refer to nodes. The networks shown here are for illustration. In existing systems fat-tree networks are realised using switches with a larger number of ports and torus networks have higher-dimensions.*

munication patters where large amounts of data is exchanged. In this case all nodes communicate with all other nodes and therefore all nodes of one half of the system have to communicate with all nodes of the other half of the system.

A fat-tree topology (see Fig. 6, left) is frequently used for Infiniband-based networks in small and medium size HPC systems. The leave nodes of the tree are the compute nodes, the other nodes are switches. While there is typically only a single path connecting a compute node to the nearest switch, there are usually multiple data paths connecting switches in order to increase the bandwidth at the higher levels, i.e. to fatten the tree. The network diameter as a function of the number of nodes N_{node} scales $\propto \log(N_{\mathrm{node}})$. The network topology is suitable for those collective communication operations which are typically implemented using binary trees, for instance global sums. For fat-tree networks it is difficult to achieve a high bi-section bandwidth when the number of nodes becomes large. For a fixed number of upstream and downstream ports per switch the bi-section bandwidth is independent of the number of nodes.
On high-end systems often d-dimensional torus networks are used with $3 \leq d \leq 5$ (see Fig. 6, right). Compared to fat-tree networks the network diameter may become large for large number of nodes as it scales $\propto N_{\mathrm{node}}^{1/d}$. Note, however, that this scaling argument only holds for very large N_{node}, while for realistic system sizes and large d, latencies are often sufficiently small. The same argument also holds for collective communication patterns, where time needed to perform, e.g., a global sum in the worst case scales $\propto d \cdot N_{\mathrm{node}}^{1/d}$. For bi-section bandwidth a good scaling $\propto N_{\mathrm{node}}^{(d-1)/d}$ is found, i.e. a scaling close to linear for high-dimensional torus networks.
However, for large d the number of network ports per node and total number of network cables becomes large. Furthermore, the distance between nodes, which within the network are nearest-neighbours, may become physically very large for big systems. As this has significant impact on the costs alternatives are considered. For instance, the TOFU network [Ajima et al.(2011)] with its 3-dimensional multi-rail torus structure features a smaller diameter and higher bi-section bandwidth than a standard 3-dimensional network with the same number of nodes. A completely different approach is taken in case of dragonfly topologies (see, e.g., [Faanes et al.(2012a)]). Dragonfly networks are multi-level networks providing an all-to-all connectivity at each level. This network topology features very small network diameters even for very large number of nodes at the expense of bi-section bandwidth scalability.
For any topology and very large number of nodes the time needed for collective communication

operations like broadcast, reduction operations as well as scatter and gather operations may have significant impact on the overall performance of the application. For this reason network designers started to provide hardware support for such communication patterns. For instance, integration of floating-point pipelines in network devices allows to avoid moving data to the processor cores just for performing a small number of arithmetic operations.

8 Power and Energy Efficiency Challenge

While compute performance continues to increase also power consumption is increasing. The power consumed by today's high-end HPC systems already exceeds 10 MWatt (see, e.g., machines listed in Table 3). As a consequence operational costs start to approach the limit of what is considered to be affordable, i.e. 20-30 MWatt. For this reason power and, more precisely, energy efficiency has become a very important topic in high-performance computing. This means that not only time-to-solution Δt but also energy-to-solution ΔE becomes relevant.
To keep operational costs in terms of expenses for electricity low, we would like to minimize energy-to-solution for a given workload W, i.e. the energy efficiency

$$\epsilon_{\mathrm{E}} = \frac{W}{\Delta E} \tag{11}$$

should be maximized. It is important to distinguish energy and power efficiency

$$\epsilon_{\mathrm{P}} = \frac{dW/dt}{P}. \tag{12}$$

For technical reasons and to keep costs of a system low there is an upper limit for the maximum power $P_{\max}$ consumed by a system (or individual system components). Energy-to-solution ΔE, time-to-solution $\Delta t = t_{\mathrm{end}} - t_{\mathrm{start}}$ and power consumption P are related by the equation

$$\Delta E = \int_{t_{\mathrm{start}}}^{t_{\mathrm{end}}} P(t)\, dt \leq P_{\max} \Delta t. \tag{13}$$

High power efficiency means that high compute performance can be provided within a small power envelope. High compute performance may result in small time-to-solution. Therefore, energy-to-solution may be small even if power consumption is high.
If the application is dominated by floating-point operations it becomes natural to quantify the workload in terms of number of floating-point operations N_{fp}. If power consumption is roughly constant we find

$$\epsilon_{\mathrm{E}} = \frac{N_{\mathrm{fp}}}{\Delta E} \simeq \frac{b_{\mathrm{fp}}\, \Delta t}{P\, \Delta t} = \epsilon_{\mathrm{P}}, \tag{14}$$

where $b_{\mathrm{fp}} = N_{\mathrm{fp}}/\Delta t$ is the average throughput of floating-point operations.
One way of comparing power efficiency has been established by the Green500 project.[11] This project ranks the most powerful HPC systems, i.e. systems which are also listed by the Top500 project, in terms of power efficiency $\epsilon_{\mathrm{P}} = b_{\mathrm{fp}}/P$ while executing the HPL benchmark. As of November 2013 an efficiency $\epsilon_{\mathrm{P}} = 4.5$ GFlop/s per Watt could be reached.[12]

[11] `http://www.green500.org`

[12] Note that the HPL benchmark does not foresee W to be kept fixed. System providers are rather allowed to tune W to optimize the resulting performance and power efficiency.

A focus on power efficiency in terms of floating-point operation throughput per Watt instead of energy efficiency can, however, lead to wrong conclusions. In [Bekas and Curioni(2010)] different algorithms for solving a linear system using Iterative Refinement methods have been investigated on a Blue Gene/P system. For a Cholesky-based version a high power efficiency was observed while for a version based on the Conjugate Gradient algorithm time- and energy-to-solution were found to be 5 and 150 times smaller, respectively.
The main reason for the huge difference in terms of energy-to-solution is the difference in energy costs for data processing on the one hand and data transport on the other hand. With today's technology the execution of a double-precision fused multiply-add operations costs about 100 pJ [Shalf et al.(2011)]. For today's commonly used SDRAM technology the average costs for reading and writing from and to external main memory are about 10 pJ/bit [Vogelsang(2010)]. Therefore reading all 3 input operands and writing 1 result costs about 2500 pJ. This means that in this example data transport is 25 times more expensive in terms of energy than data processing. This observation is not expected to change in the future. The use of algorithms with higher data locality, as was shown in [Bekas and Curioni(2010)] for a specific example, will become critical as energy efficiency becomes more important.

9 Continuous Increase of Parallelism

In the previous section we have seen that energy efficiency has become a major concern for high-performance computing. Increasing core clock speed tends to lead to reduced energy efficiency. It furthermore makes power distribution and cooling more challenging and thus more expensive. Further increase in performance thus requires further increase in parallelism. In this lecture we encountered the following levels of parallelism:

Instruction-level parallelism (ILP): Compute devices comprise multiple processing pipelines where instructions can be processed in parallel even if the architecture appears to be scalar. It is mainly task of the compiler to exploit ILP.

Data-level parallelism (DLP): DLP is exploited by SIMD instructions where the same instructions are applied to a parallel stream of data. Compilers may be able to detect this parallelism and generate suitable instructions, alternatively application programmers have to explicitly leverage extensions to standard programming languages to enforce the use of SIMD instructions.

Thread-level parallelism (TLP): To use multiple processor cores or, in case of SMT, to better utilize multiple processing pipelines, computing devices may support concurrent execution of multiple threads of instructions. On processors POSIX threads or OpenMP are established programming models for multi-threaded programming.

Process-level parallelism: To further increase parallelism it becomes necessary (or more efficient) to start multiple processes which process data within a private address space. The underlying memory may either be still shared by multiple processes or physically disjunct, e.g. reside on different nodes. Data has in both cases to be exchanged explicitly by sending and receiving messages. The most widely used communication interface is the Message Passing Interface (MPI) (see [MPI Forum(2012)] for the latest version of this standard). It is based on a model where data is moved from the address space of one process to that of another process through cooperative operations on each process.

MPI provides a full abstraction of the underlying hardware architecture which makes applications using MPI very portable. Details of the underlying network architecture do, however, effect the performance. The use of MPI will be topic of a later lecture.

When parallelizing applications over N computing devices (which here may refer to multiple processing pipelines, cores or nodes), one would like to keep the parallel speedup $S(N)$ close to the N. The *parallel speedup* is defined as

$$S(N) = \frac{\Delta t(1)}{\Delta t(N)}, \tag{15}$$

where $\Delta t(N)$ is time-to-solution using N computational devices. It is convenient to consider the *parallel efficiency* which is the parallel speedup weighted by a factor $1/N$ to account for the N-fold increase of computational resources:

$$\epsilon_{\rm p} = \frac{S(N)}{N}. \tag{16}$$

As parallel efficiency $\epsilon_{\rm p}$ is crucial for scaling to a very large number of nodes it is important to understand the limiting factors. Even if parallelization overheads could be completely removed, $\epsilon_{\rm p}$ might be small due to the fraction of a task which cannot be parallelized. Let us assume that a fraction $f > 0$ of a given task can be parallelized. The serial execution time can then be written as $\Delta t(1) = \Delta t_s + \Delta t_p$, where $\Delta t_s = (1-f)\Delta t$ and $\Delta t_p = f\Delta t$ refer to the execution time of the part of the task which cannot and can be executed in parallel, respectively. Neglecting any overheads due to parallelization, Δt_p reduces by a factor N when executed on N processing devices in parallel. With this assumption we find $\Delta t(N) = \Delta t_s + \Delta t_p/N$ and therefore parallel speedup becomes

$$S(N) = \frac{(1-f)\Delta t + f\Delta t}{(1-f)\Delta t + f\Delta t/N} = \frac{1}{(1-f) + f/N}. \tag{17}$$

This model analysis is called *Amdahl's law* [Amdahl(1967)]. There is always a fraction of a computational task which cannot be parallelized, i.e. $f < 1$. From Amdahl's law it follows that parallel speedup is limited by $\lim_{N\to\infty} S(N) = (1-f)^{-1}$. Even if a large fraction of a task, which, say, accounts for 90 % of the execution time when executed in serial mode, can be perfectly parallelized, the maximum speedup is 10.

Amdahl's law has been criticized by Gustafson [Gustafson(1988)] as being unrealistic because it assumes a fixed workload. From Amdahl's law it follows how parallel speedup scales with fixed workload, i.e. the so-called *strong scaling* case is considered. Alternatively one may consider *weak scaling* where the workload changes linearly with the number of computational devices. In case of perfect weak scaling, i.e. in the absence of any parallelization overheads, the time Δt_p spent in the parallel section of a computational task would stay constant. Concerning the serial part Gustafson argues, based on empirical observations, that this part is typically independent of the workload.[13] The following scaled parallel speedup is known as *Gustafson's law*:

$$\tilde{S}(N) = \frac{(1-f)\Delta t + Nf\Delta t}{(1-f)\Delta t + f\Delta t} = N - (N-1)(1-f). \tag{18}$$

From Gustafson's law a scaled parallel efficiency $\tilde{S}(N)/N > f$ can be derived. Unlike in the strong scaling case described by Amdahl's law, scaled parallel efficiency does not approach zero for large N.

[13] This empirical basis for this assertion was not very strong and in many numerical tasks this does not hold.

	Sequoia	K Computer	Piz Daint
System architecture	Blue Gene/Q	K Computer	XC30
Vendor	IBM	Fujitsu	CRAY
Top500 (11/2013)	#3	#4	#6
Processor type	Blue Gene/Q A2	Sparc64 VIIIfx	Xeon E5-2650
N_{core}	1 572 864	663 552	42 176
Accelerator type	–	–	K20 GPU
N_{acc}	–	–	5 272
Network topology	5d torus	3d toroidal	dragonfly
Link bandwidth B_{link} [GByte/s]	2	5	4.7-5.25
Floating-point peak B_{fp} [PFlop/s]	20.1	10.6	7.8
Memory capacity C_{mem} [PiByte]	1.5	1.3	0.2
Power [MWatt]	7.9	12.7	2.3

Table 3: *Hardware parameters of selected high-end HPC systems: Blue Gene/Q [Aho et al.(2013)] installation at LLNL (USA), K Computer [Yokokawa(2012)] at RIKEN (Japan), Piz Daint at CSCS (Switzerland). Power refers to the amount of electricity consumed during execution of the HPL benchmark as reported in the Top500 list of November 2013.*

10 Today's Supercomputer Architectures

In this section we discuss a selected set of parallel architectures and existing incarnations which as of today are used for large-scale scientific applications.

IBM Blue Gene/Q [Aho et al.(2013)] is the third generation of a series of architectures which first was introduced in 2004. A unique feature of this architecture is the integration of the network on the processor. In this way very high bandwidth to the network and extremely small latencies can be achieved, which has a strong impact on the scalability properties of the architecture. As can be seen from Table 1 the number of cores per processor is large. To efficiently utilize the processing pipelines up to 4 hardware threads may have to be started resulting in up to 64 threads per processor. A separate, 17th core is used just for running a slim operating system to minimize interference of computational task execution and asynchronous management of system services. The nodes are interconnected in a 5-dimensional torus using 10 out of 11 available network ports per processor. Since the Blue Gene/Q architecture has become generally available in 2012 O(10) systems have been installed, including an installation at Jülich Supercomputing Centre, called JUQUEEN, comprising 458 752 cores. In Table 3 we show selected hardware parameters of the largest existing Blue Gene/Q installation at Lawrence Livermore National Laboratory (LLNL, USA).

Also the K-Computer [Yokokawa(2012)] developed by Fujitsu together with the Japanese research lab RIKEN is based on a processor which has been specifically designed for high-end HPC systems. Details of the processor and the only existing installation of the architecture are listed in Table 1 and 3, respectively. The processors are connected to a custom network chip which is part of the TOFU network [Ajima et al.(2011)] (see section 7 for more details). Compared to Blue Gene/Q the architecture provides relatively complex processor cores with more features, higher memory bandwidth and larger memory capacity. This makes the system easier

to use for a larger range of applications. The price is a significantly lower power efficiency.
The Cray XC30 system [Faanes et al.(2012b)] is unlike the other systems based on a commodity processor. These are interconnected by a special high-performance network using a dragonfly topology. The parameters of a Cray XC30 system at the Swiss supercomputing centre CSCS can be found in Table 3. The number of nodes is relatively small as each of the nodes provides significantly more compute performance thanks to GPU-based compute accelerators. Because of these accelerators the power consumption when executing the HPL benchmark is significantly smaller compared to the two other systems.

11 Outlook

As of today parallel computer architectures are capable of providing $O(10)$ PFlop/s compute performance. Around 2017 and after 2020 it is expected that with next generation of parallel computer architectures systems will become available with a peak compute performance of $O(100)$ PFlop/s and $O(1)$ EFlop/s, respectively. This will, however, only be possible by yet another significant increase of parallelism. With $O(10^7)$ threads on today's machines the amount of concurrency which has to be exposed by the application is already very high. It will require significant efforts from developers of algorithms and applications to leverage the increased performance of future parallel computer architectures. While challenging in use, high-end HPC systems open significant opportunities for scientific research and the creation of new scientific knowledge and results.
To learn more about computer architectures in general and the use of high-performance computing systems for scientific applications we recommend [Hennessy and Patterson(2011)] and [Hager and Wellein(2010)] for further reading.

Acknowledgements

The author would like to thank Hubert Simma (DESY, Germany) for many valuable discussions and ideas which found their way into these lecture notes but have not been published elsewhere.

References

[Neumann(1945)] J. v. Neumann, Tech. Rep. (1945).

[Moore(2006)] G. E. Moore, Solid-State Circuits Society Newsletter, IEEE **11**, 33 (2006), ISSN 1098-4232.

[Bohr(2007)] M. Bohr, Solid-State Circuits Society Newsletter, IEEE **12**, 11 (2007), ISSN 1098-4232.

[Wienke et al.(2013)] S. Wienke et al., in *Supercomputing*, edited by J. Kunkel, T. Ludwig, and H. Meuer (Springer Berlin Heidelberg, 2013), vol. 7905 of *Lecture Notes in Computer Science*, pp. 330–342, ISBN 978-3-642-38749-4, URL http://dx.doi.org/10.1007/978-3-642-38750-0_25.

[Terpstra et al.(2010)] D. Terpstra et al., in *Tools for High Performance Computing 2009*, edited by M. S. Mller, M. M. Resch, A. Schulz, and W. E. Nagel (Springer Berlin Heidelberg, 2010), pp. 157–173, ISBN 978-3-642-11260-7, URL http://dx.doi.org/10.1007/978-3-642-11261-4_11.

[Bilardi et al.(2005)] G. Bilardi et al., in *Proceedings of the 12th International Conference on High Performance Computing* (Springer-Verlag, Berlin, Heidelberg, 2005), HiPC'05, pp. 386–397, ISBN 3-540-30936-5, 978-3-540-30936-9, URL http://dx.doi.org/10.1007/11602569_41.

[Culler et al.(1993)] D. Culler et al., in *Proceedings of the Fourth ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming* (ACM, New York, NY, USA, 1993), PPOPP '93, pp. 1–12, ISBN 0-89791-589-5, URL http://doi.acm.org/10.1145/155332.155333.

[Harris(2005)] M. Harris, in *ACM SIGGRAPH 2005 Courses* (ACM, New York, NY, USA, 2005), SIGGRAPH '05, URL http://doi.acm.org/10.1145/1198555.1198768.

[Williams et al.(2009)] S. Williams et al., Commun. ACM **52**, 65 (2009), ISSN 0001-0782, URL http://doi.acm.org/10.1145/1498765.1498785.

[Intel(2013)] Intel, *ARK product specifications* (2013), URL http://ark.intel.com.

[Aho et al.(2013)] M. Aho et al., IBM Journal of Research and Development **57**, 0:1 (2013), ISSN 0018-8646.

[Maruyama et al.(2010)] T. Maruyama et al., IEEE Micro **30**, 30 (2010), ISSN 0272-1732, URL http://dx.doi.org/10.1109/MM.2010.40.

[Kowarschik and Weiß(2003)] M. Kowarschik and C. Weiß, in *Algorithms for Memory Hierarchies*, edited by U. Meyer, P. Sanders, and J. Sibeyn (Springer Berlin Heidelberg, 2003), vol. 2625 of *Lecture Notes in Computer Science*, pp. 213–232, ISBN 978-3-540-00883-5, URL http://dx.doi.org/10.1007/3-540-36574-5_10.

[Hager and Wellein(2010)] G. Hager and G. Wellein, *Introduction to High Performance Computing for Scientists and Engineers* (CRC Press, Inc., Boca Raton, FL, USA, 2010), 1st ed., ISBN 143981192X, 9781439811924.

[Drepper(2007)] U. Drepper, *What every programmer should know about memory* (2007).

[Ajima et al.(2011)] Y. Ajima et al., in *High Performance Interconnects (HOTI), 2011 IEEE 19th Annual Symposium on* (2011), pp. 87–94.

[Faanes et al.(2012a)] G. Faanes et al., in *High Performance Computing, Networking, Storage and Analysis (SC), 2012 International Conference for* (2012a), pp. 1–9, ISSN 2167-4329.

[Bekas and Curioni(2010)] C. Bekas and A. Curioni, Computer Science - Research and Development **25**, 187 (2010), ISSN 1865-2034, URL http://dx.doi.org/10.1007/s00450-010-0119-z.

[Shalf et al.(2011)] J. Shalf et al., in *High Performance Computing for Computational Science – VECPAR 2010*, edited by J. M. L. M. Palma, M. Daydé, O. Marques, and J. C. Lopes (Springer Berlin Heidelberg, 2011), vol. 6449 of *Lecture Notes in Computer Science*, pp. 1–25, ISBN 978-3-642-19327-9, URL http://dx.doi.org/10.1007/978-3-642-19328-6_1.

[Vogelsang(2010)] T. Vogelsang, in *Microarchitecture (MICRO), 2010 43rd Annual IEEE/ACM International Symposium on* (2010), pp. 363–374, ISSN 1072-4451.

[MPI Forum(2012)] MPI Forum, *MPI: A message-passing interface standard version 3.0* (2012), URL http://www.mpi-forum.org/docs/mpi-3.0.

[Amdahl(1967)] G. M. Amdahl, in *Proceedings of the April 18-20, 1967, Spring Joint Computer Conference* (ACM, New York, NY, USA, 1967), AFIPS '67 (Spring), pp. 483–485, URL http://doi.acm.org/10.1145/1465482.1465560.

[Gustafson(1988)] J. L. Gustafson, Commun. ACM **31**, 532 (1988), ISSN 0001-0782, URL http://doi.acm.org/10.1145/42411.42415.

[Yokokawa(2012)] M. Yokokawa, in *Networking and Computing (ICNC), 2012 Third International Conference on* (2012), pp. 21–22.

[Faanes et al.(2012b)] G. Faanes et al., in *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis* (IEEE Computer Society Press, Los Alamitos, CA, USA, 2012b), SC '12, pp. 103:1–103:9, ISBN 978-1-4673-0804-5, URL http://dl.acm.org/citation.cfm?id=2388996.2389136.

[Hennessy and Patterson(2011)] J. L. Hennessy and D. A. Patterson, *Computer Architecture, Fifth Edition: A Quantitative Approach* (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2011), 5th ed., ISBN 012383872X, 9780123838728.

D 2 Parallel Programming[1]

A. Schnurpfeil, F. Janetzko
Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Even though the number of transistors on integrated circuits doubles every 12 to 24 months according to Moore's law resulting in high performing CPUs, the sequential execution of code on such computer chips is by far not sufficient to meet the requirements of a wide range of applications, in particular the performance needs of scientific simulations extend far beyond this level. Nowadays, it is necessary to have systems containing a large number of CPUs and therefore, more and more complex and powerful supercomputers are being installed to meet these requirements. The *top500 list* [1] gives an impressive overview of the state-of-the-art parallel supercomputers which provide a theoretical peak performance of up to 55 PFlops, i.e. $55 * 10^{15}$ floating point operations per second. However, this overwhelming computing power sets special requirements on the hardware architectures which in turn tremendously affect the software requirements in order to exploit the potential of such supercomputers. Therefore, the current situation essentially puts two demands on the user of such systems, on the one hand a general understanding of the underlying hardware and on the other hand special knowledge about the possibilities to produce software that benefits from the hardware. In section 2 subsection 2.1 focuses on the former part and gives an introduction to the basic concepts on which all computer systems are based these days. It explains the different use cases which result thereof and it describes the major memory models. The remaining parts of section 2 describe the software, starting with parallel programming models in subsection 2.2, with special focus on implications which follow from distributed memory and shared memory systems, respectively. In addition to that some design strategies are discussed in subsection 2.3 to cope with the requirement to balance the work load of parallel programs on multiprocessor architectures. Finally, subsection 2.4 discusses some concepts to understand and to describe the performance potential of parallel programs. The article proceeds with two programming models commonly used these days, namely *Message Passing Interface* (MPI) [2] for distributed memory architectures in section 3 and *Open Multi-Processing* (OpenMP) [3] for shared memory architectures in section 4. After the introduction of MPI and OpenMP, this article ends with an appendix showing a simple hybrid code example, i.e. both concepts are applied concurrently to exploit the potential of hybrid distributed-shared memory systems. In fact, this approach becomes more and more necessary to get the most out of a wide range of today's architectures.

2 Concepts and Models of Parallel Computing

2.1 Hardware Terminology and Concepts

John von Neumann developed more than half a century ago a concept for a universal computer architecture [4], named after him *von-Neumann Architecture* (Fig. 1). Since then basically all electronic computers followed this design and still nowadays all modern computer systems contain the four main components: (i) *Memory*, usually random access memory (RAM), which is used to store instructions and data, (ii) the *Control Unit* coordinates the data and instruction streams by fetching, decoding and determining the sequence in which data and instructions are processed, (iii) the *Arithmetic Logic Unit* (ALU), which performs arithmetic operations on the data, and finally (iv) the *Input/Output* interface. The Control Unit and the ALU are parts of the *Central Processing Unit* (CPU), sometimes simply called *processor*.

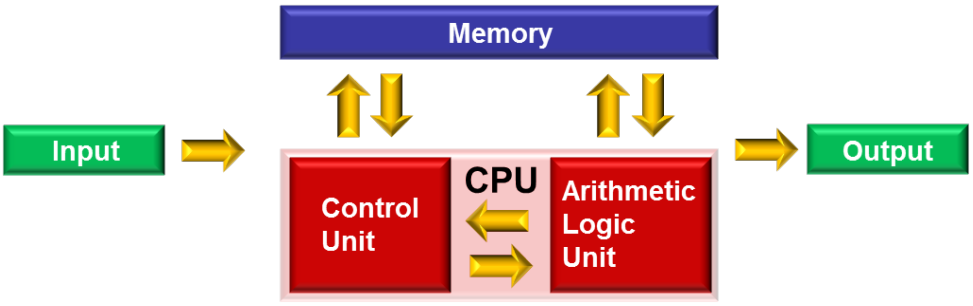


Fig. 1: *Basic concept of all modern computer systems: The von-Neumann-Architecture for electronic computers.*

Parallel computer systems can be classified according to the parallelism they show with respect to the instruction and data streams. A commonly used classification is *Flynn's Taxonomy* [5] which divides parallel computers into four categories:

SISD, *Singe Instruction Single Data*. In this category fall all all computers which process only one instruction and one data stream at a time. An example of such a CPU was the first Intel Pentium processor.

SIMD, *Single Instruction Multiple Data*. Here a single instruction stream is used for several data streams at the same time. One example are operations on vectors like $\mathbf{c} = \mathbf{a} + \mathbf{b}$, where the instruction '+' is simultaneously applied to different elements of the vectors. This kind of parallelization is therefore often called *vectorization*. Classical vector computers as well as most modern CPUs (for example, x86 processors with *Streaming SIMD Extensions*, SSE) belong to this category.

MISD, *Multiple Instruction Single Data*. This refers to an architecture, where different instructions are applied to one data stream. Not in common use, only few systems based on this approach have ever been built.

MIMD, *Multiple Instruction Multiple Data*. Multiple instruction streams are applied to multiple data streams. This approach contains the SIMD approach and all modern *multiprocessor architectures* belong to this category.

In multiprocessor architectures several CPUs are integrated into one computer system. The smallest compute part of such a system is called a *compute node* or simply a *node*. It contains usually one or more CPUs, memory and an interface to the network, which connects the nodes. Furthermore, modern CPUs are equipped with a hierarchical *cache* structure. Data stored in the cache can be accessed much faster than data that resides in memory.
If all CPUs share a single physical memory address space the architecture is called a *shared memory* architecture. Such a system is shown schematically in Fig. 2.

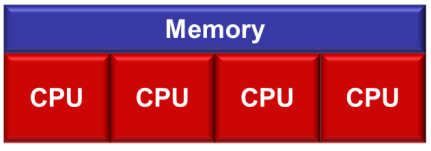


Fig. 2: *A shared memory architecture (schematically), where all CPUs share a single physical address space.*

All CPUs have the same access time to the shared (global) memory. Such systems are called *uniform memory access* (UMA) systems, sometimes also known as *Symmetrical Multi Processors* (SMP). If the hardware takes care about the *cache coherency*, which means that if one CPU updates a memory location, caches of all CPUs which contain this address will be updated as well, then such a system has a *cache-coherent uniform memory access* (cc-UMA) architecture. Programming such architectures is straightforward due to the availability of a global address space which makes it very easy to share data between the CPUs. However, the user has to ensure that the memory accesses are synchronized properly in order to guarantee the correctness of the data.
A multiprocessor architecture where each CPU has its own memory with its own address space is called a *distributed memory* architecture (Fig. 3).

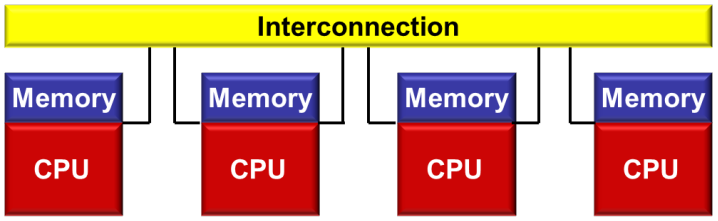


Fig. 3: *Scheme of a distributed memory architecture. Each CPU has its own memory and address space.*

The CPUs are operating independently on their local memory and therefore, the concept of cache coherency is not applicable. If a CPU needs to access data that resides in the memory of another CPU, explicit passing of this data via the interconnect is necessary. This is a *Non-Uniform Memory Access* (NUMA) architecture, because access times to local and remote memory differ. *Remote Memory Access* (RMA) architectures have dedicated hardware support for accessing the memory of a remote node in contrast to *No Remote Memory Access* (NORMA) architectures. To program distributed memory architectures the user needs to adapt algorithms and especially to distribute data structures efficiently. He also must take care of the communication between the compute nodes explicitly.
Most modern multiprocessor architectures are *hybrid distributed-shared memory* architectures, having shared-memory nodes with one or more SMPs per compute node and (nowadays in High Performance Computing systems) up to several hundred thousand nodes. The SMPs can consist of, for example, CPUs with several cores and can be additionally equipped with *Graphics Processing Units* (GPUs, often simply called graphics cards) or other multi- or many-core devices, like the Intel®MIC (Many Integrated Core) architecture. A hybrid distributed-shared memory system with two CPUs per node is depicted in Fig. 4.

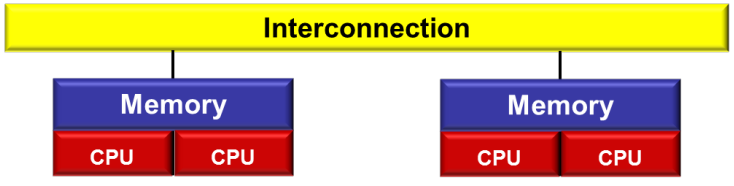


Fig. 4: *Scheme of a hybrid distributed-shared memory architecture with two CPUs per compute node.*

The system has shared memory nodes. But since the memory per node is divided into memory banks which are connected to different CPUs the memory access times within one compute node are not all equal (even though there is a single address space per node). In general cache coherency is guaranteed in such systems and therefore the nodes have a cc-NUMA architecture and form together a distributed memory system.

2.2 Parallel Programming Models and Concepts

In order to benefit from parallel computers the work to be done has to be distributed over processors. In principle there are two approaches how to partition the work [6, 7].

Computation-centric approach. In this approach different instructions on the data are performed on different processes (also called *functional decomposition*). For example, for a set of values **y** the minimum, the maximum and the average has to be calculated (Eq. 1).

$$\begin{aligned} y_{\min} &= \min(\mathbf{y}) \\ y_{\max} &= \max(\mathbf{y}) \\ y_{\text{avg}} &= \text{avg}(\mathbf{y}) \end{aligned} \tag{1}$$

This work can be distributed in such a way that all three operations are performed on **y** simultaneously, for example each on a different processor.

Data-centric approach. In the data-centric approach the data to be processed is split and distributed over the compute nodes (*domain decomposition*). Suppose the operations in Eq. 2 have to be performed.

$$\mathbf{z} = \alpha\mathbf{y} + \mathbf{x} \tag{2}$$

In the data-centric approach the elements of **y** and **x** are distributed over the available processors so that all processors perform the instructions simultaneously but only for parts of the vectors. After the computation is done the results are collected into one result vector.

Frequently, these two approaches are combined. For example, for a large vector **y** each of the operations in Eq. 1 can be domain decomposed by distributing **y** over several processors.
Analogical to the terms SIMD and MIMD, which refer to hardware (see section 2.1), the *Single Program Multiple Data* (SPMD) and the *Multiple Program Multiple Data* (MPMD) programming paradigms refer to two parallel software models. In the SPMD paradigm the same program (executable) is executed on all participating CPUs. It is the programmer's responsibility to split and distribute the work (i.e. instructions, data or both) among the CPUs, for example using conditional constructs like *if* or *select* statements. In the MPMD paradigm different executables are used on the participating CPUs, which can be beneficial for algorithms that fit better to the computation-centric partition approach than to the data-centric one.
Depending on the way how different programs or different instances of a program interact with each other one distinguishes two parallel programming interaction models.

Message Passing Model. In this model a set of parallel processes is executed, each with its private data structures. Exchange of data between the processes is only possible by explicitly passing messages from one process to the other. This model is designed to run on distributed memory systems, where no global address space exists. However, it can be used also on shared memory systems. An example implementation of this model is the very commonly used *Message Passing Interface* (MPI) [2], which will be discussed in more detail in Section 3.

Shared Memory Model. It is based on a number of threads with a common global address space, which are connected to one process. In most applications the parallelism is restricted in this model to code regions where operations can be executed in parallel, creating parallel loops or parallel sections. Threads that execute the same parallel region are called a *team*. This type of execution is also called *fork-join parallelism*. An advantage of this model is that automatic parallelization is supported by most compilers. While this leads to a first working parallel program, usually manual refinements are necessary in order to improve the efficiency. A disadvantage of the shared memory model is that it and can be used only on shared memory architectures. A widely used implementation of this model is *OpenMP*, discussed in Section 4.

Sometimes a third model is described which is called *Implicit Model*. In this model the interaction between the different instances (processes or tasks) is not visible to the programmer. It is just necessary to specify which data is needed and the compiler will take care of retrieving it to the appropriate location. This model is implemented in the various *Partitioned Global Address Space* (PGAS) languages like *Co-Array Fortran* (CAF) and *Unified Parallel C* (UPC).

2.3 Design Strategy for Parallel Programs

Designing efficient parallel programs can be a major challenge. Following the ideas of Ian Foster [6] and Michael J. Quinn [7] we will discuss here briefly a design strategy for parallel programs based on the task/channel model.
In the task/channel model a process, its local memory and a collection of input and output ports is called a *task* [7]. Tasks can communicate with each other via *channels* (non-local communication), where each channel connects the output port of one task with the input port of another task. *Primitive tasks* are the smallest logical set of instructions and algorithm can be split in.
According to Foster [6] the process of designing a parallel program can be split into four parts: (i) *Partitioning*, (ii) *Communication*, (iii) *Agglomeration*, and (iv) *Mapping* as show in Fig. 5.

Partitioning. In this first step the problem which should be solved is partitioned into primitive tasks. A functional decomposition, a domain decomposition or a combination of both approaches can be used as discussed in section 2.2. Already at this stage it is useful to have an idea about the hardware the program will run on, since a good partition scheme should have a ratio of primitive tasks to number of processors (or cores) of at least 10:1. Furthermore, redundant computations and redundant storage of data should be avoided, the primitive tasks should be of comparable size and the number of primitive tasks should scale with the *problem size*.

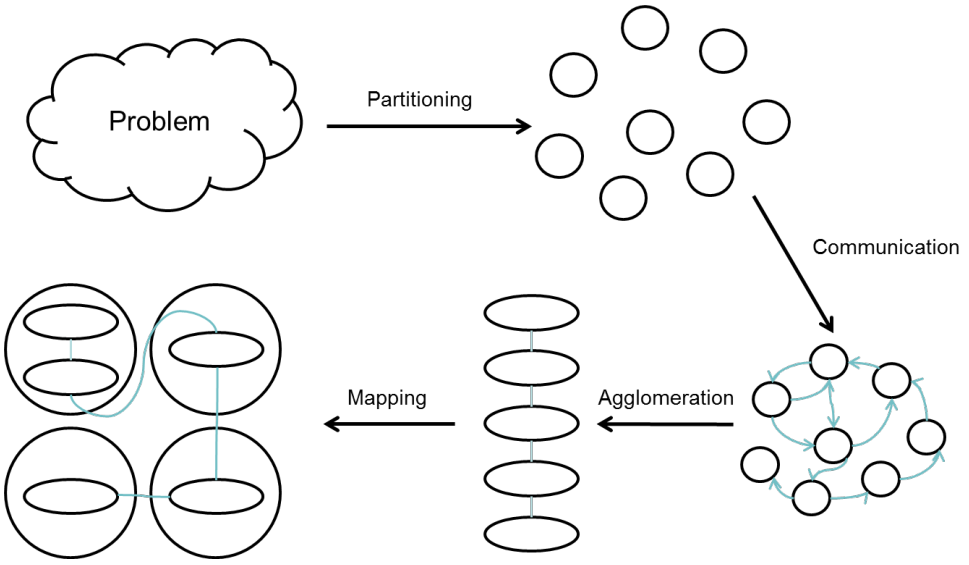


Fig. 5: *Foster's Design Strategy consisting of the four steps Partitioning, Communication, Agglomeration, and Mapping.*

Communication. Once the problem is split into primitive tasks, the communication between them needs to be analyzed. Two types of communication can be identified. *Local communication* means that a (primitive) task needs data from no or only a small number of other (primitive) tasks while *Global Communication* occurs when all (primitive) tasks are involved in a communication step. In order to minimize the parallel overhead which is introduced by any kind of communication it should be kept to the minimum and one should check that the communication operations are balanced among the (primitive) tasks. Furthermore, communication should occur concurrently.

Agglomeration. The aim of this step is to move from the more abstract first two steps towards a more concrete concept how the problem can be solved on a certain hardware architecture. This is achieved by analyzing the primitive tasks and communication patterns obtained in the previous steps and reducing the number of primitive tasks by grouping them together to actual tasks which are going to be executed by processors of the target architecture. A good guideline here is to increase the locality of the communication. For example, frequent communication between two primitive tasks can be eliminated by grouping them together into one task. Furthermore, it might be worth to replicate some computations if this avoids more time-consuming communication. Again, the communication and computation per task (*load balancing*) should be balanced. Finally, the number of tasks obtained should fit for the given problem size to the target computer architecture and the number of tasks should be an increasing function of the problem size.

Mapping. The last step actually assigns the tasks to the processors of the parallel computer system by coding the algorithm into a program. Here, one is sometimes confronted with two conflicting goals. On the one hand one wants to maximize the system utilization, i.e. one wants to use as many processors as possible. On the other hand the inter-processor communication should be kept as low as possible.

To find a compromise between both goals is one of the aims in this step. Also the actual parallelization model needs to be chosen. Whether to chose a pure message passing model, a pure shared memory model, or a combination of both models and which implementation of these models to use, depends on the algorithm to be parallelized and on the target computer architecture. In general, a hybrid MPI/OpenMP parallelization scheme offers the best potential to make efficient use of modern HPC architectures.

2.4 Efficiency and Scalability of Parallel Programs

In order to check the quality of the parallelization scheme for a certain problem or algorithm, i.e. how efficiently a parallel program makes use of the given hardware, the performance and efficiency of a code needs to be analyzed and quantified. For an in-depth analysis of codes various performance analysis tools exist, whose detailed discussion is however beyond the scope of this article. However, some basic terms and concepts for quantifying the quality of a parallel program are discussed in the following.

Speedup. The speedup of a program or algorithm shows how much faster the parallel version runs on n processors compared with the serial version. The mathematical definition of the speedup $S(n)$ is given in Eq. 3.

$$S(n) = \frac{t(1)}{t(n)} \tag{3}$$

Here, $t(1)$ refers to the execution time of the serial code and $t(n)$ to the execution time of the parallel version on n processors. A speedup of > 1 means the parallel codes runs faster than the serial one and a *linear scaling*, which means doubling n halves the execution time $t(n)$, indicates a perfect scalability of the application.

Efficiency. While the speedup tells only whether the parallel code runs faster than the serial one the efficiency gives a more quantitative picture how well the code performs in parallel. The efficiency $E(n)$ on n processors is defined in Eq. 4.

$$E(n) = \frac{t(1)}{n \cdot t(n)} \cdot 100\% \tag{4}$$

It denotes the percentage of the ideal speedup that was actually achieved, where $t(1)$ refers again to the execution time of the serial code and $t(n)$ to the execution time of the parallel version on n processors.

Scalability. The scalability shows how a code behaves when it is executed on an increasing number of processors. In general, two types of scalability are distinguished: (i) *weak scaling*, the problem size is increased proportionally to the number of processors and (ii) *strong scaling*, the problem size is kept fixed while increasing the number of processors. In weak scaling the number of operations per processor stays constant. Therefore, for a perfect weak scaling a constant execution time and therefore a speedup of 1 on different numbers of processors is expected as shown in Fig. 6. In contrast to this the number of operations decreases in strong scaling and the ideal strong scaling shows a constantly increasing speedup (Fig. 6).

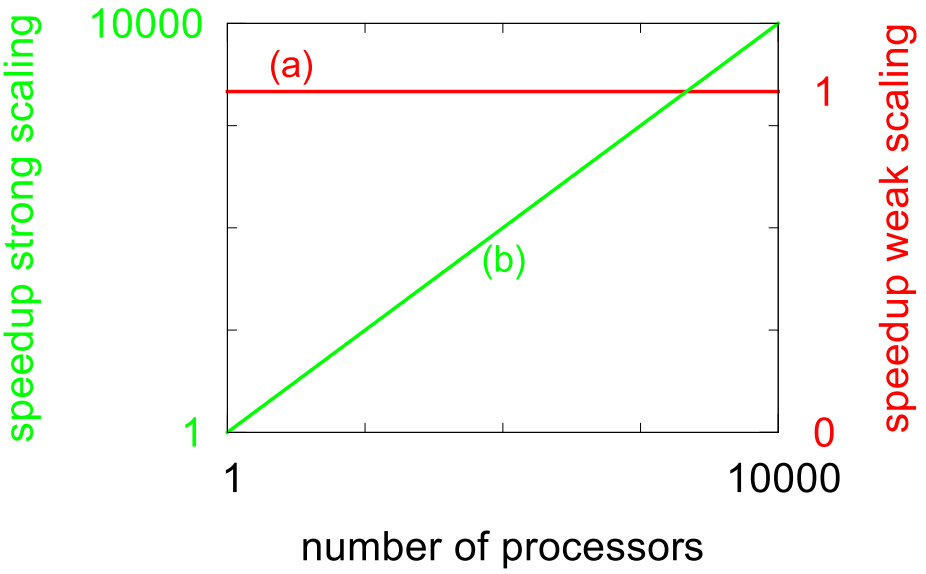


Fig. 6: *Speedup of (a) weak scaling and (b) strong scaling experiments.*

The parallel performance and efficiency of a code is limited by the parts that cannot be parallelized. In order to estimate the maximum speedup which can be expected for a certain parallel code *Amdahl's Law* [8] can be applied. According to this the real speedup S_r on n processors that can be expected from a code with a serial part α is given by Eq. 5.

$$S_r = \frac{1}{\alpha \cdot \frac{1-\alpha}{n}} \tag{5}$$

For example, if a parallel program has 10% serial code ($\alpha = 0.1$) then the real speedup on $n = 8$ processors is $S_r = 4.7$. It is also intriguing to realize that Eq. 5 converges for large n to $\lim_{n\to\infty} = 1/\alpha$. This means that the (theoretical) maximum speedup of a parallel program is solely determined by the percentage of its serial part, and it is independent of the number of processors.

3 The Message Passing Interface (MPI)

The *Message Passing Interface* (MPI) [2] is a nowadays widely-used industry standard for the message passing programming model introduced in Section 2.2. Its development started in the early 1990's and the first standard (MPI-1.0) was published 1994. Since then it has been continuously developed further and the latest version (MPI-3.0) was released in September 2012. This standard provides specifications for the message passing programming model including bindings for C, Fortran 77, Fortran 90, and Fortran 2008. It is usually implemented as a library which needs to be linked to the applications and different implementations of this standard exist. Applications that comply with the standard are portable to any computer for which an MPI implementation is available.

3.1 Terminology and Concepts

MPI implementations provide routines for the exchange of data between processes (communication) as well as for parallel I/O. The most important basic terms and concepts needed for an efficient usage of MPI are briefly discussed in the following.

Task. A Task is an instance, sub-program or process of an MPI program. It is the smallest parallel unit in MPI with its own private memory. Communication in MPI takes place between tasks, where a task can communicate with other tasks as well as with itself.

Message. A message in MPI refers to a packet of data which needs to be exchanged between tasks. A message consists of two parts: a data part, containing the actual data, specified via the memory address, the number of elements and the datatype of the elements to be sent, and a message *envelop*, where among other information source (sending task(s)) and destination (receiving task(s)) of the message are given.

Context. Communications in MPI always take place within a certain context. A context is like an additional tag that differentiates messages. This way messages that are sent within different contexts do not interfere with each other, which is especially important for building parallel libraries.

Communicator. This is the basic communication concept of MPI. An MPI communicator consists of a group of tasks and a context. The default communicator which contains all MPI tasks is referenced by the MPI object *MPI_COMM_WORLD*.

Rank. This is an unique identifier assigned to each task of an MPI program. Ranks are integer numbers starting at 0. A task can have different ranks in different communicators.

MPI routines have certain properties which determine their communication behavior. The performance of an MPI-parallelized code depends critically on the use of routines with the correct properties for proper parts of the algorithm.

Synchronous/Asynchronous. The program will return successfully from a *synchronous* routine only if the required matching operation has started (e.g. *sending - receiving*). In case of an *asynchronous* or *buffered* routine the program will return successfully immediately, i.e. whether or not the required matching operation has started or not.

Blocking/Nonblocking. A procedure is *blocking* if returning from the procedure indicates, that it is allowed to reuse the resources specified in the call to the procedure. Accordingly, returning from a nonblocking procedure resources cannot be reused until the communication has been completed by calling an appropriate procedure.

3.2 Language Bindings and Basic Routines

The MPI definitions are included using the appropriate include files. For C this is done with

```
#include <mpi.h>
```

For Fortran the syntax depends on the standard to be used:

```
include 'mpif.h'     ! Fortran 77
use mpi              ! Fortran 90
use mpi_f08          ! Fortran 2008
```

When compiling the application the proper MPI library needs to be linked with the code. Most MPI implementations offer appropriate compiler wrappers for this purpose.
All MPI functions follow a generic format. They start with the prefix *MPI_* followed by the function name and the corresponding parameters. For C an error code is returned, the Fortran routines are called with an extra parameter for the error code. The following four basic MPI routines can be found in basically all MPI programs:

MPI_Init. This is usually the first call to MPI and initializes the library. After this call all other MPI routines can be used. Other initialization routines are available for hybrid MPI/OpenMP codes.

MPI_Finalize. This is usually the last call to MPI and should be called at the end of the MPI program.

MPI_Comm_size. Determines the number of MPI tasks the code is executed with.

MPI_Comm_rank. For each task the routine returns its rank.

3.3 Types of Communication

MPI offers different types of communication:

Point-To-Point Communication. The communication between two and only two tasks is called point-to-point communication. A source task sends a message to a destination task using an **MPI_Send** routine. The destination task issues a **MPI_Recv** routine to receive the message. Every message sent with a point-to-point call must be matched by a corresponding receive call. MPI offers blocking point-to-point routines as well as nonblocking ones. The routines can be furthermore either synchronous or asynchronous.

Selected blocking point-to-point routines

MPI_Ssend (synchronous send)

MPI_Bsend (asynchronous or buffered send)

MPI_Recv (blocking receive)

Selected nonblocking point-to-point routines

MPI_Issend (synchronous send)

MPI_Ibsend (asynchronous or buffered send

MPI_Irecv (nonblocking receive)

Collective Communication. In this type of communication *all* tasks of a communicator must participate. Different routines are available. Like in the case of the point-to-point communication for most blocking collective routines corresponding nonblocking routines are available by inserting an 'I' after the *MPI_* prefix. There are three main classes of MPI collective communication routines: (i) one task sends messages to all others (*One-to-All*) (ii) one task receives messages from all other tasks (*All-to-One*) and (iii) all tasks communicate with each other(*All-to-All*). An schematic overview over the different classes for collective routines is given in Fig. 7.

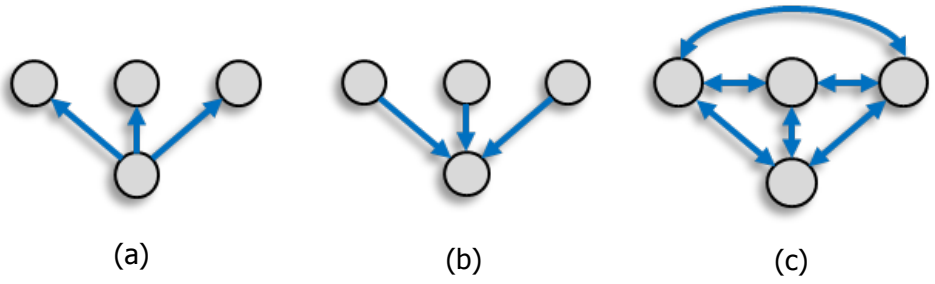


Fig. 7: *Schematic view of the three main classes of collective communication MPI routines. (a) One-to-All communication, (b) All-to-One communication, (c) All-To-All communication.*

A few selected collective operations are:

MPI_Bcast. One task which is called *root* task sends a message to all other tasks. This routine is an example for a *One-to-All* communication routine.

MPI_Reduce. This belongs to the class of *All-to-One* communication routines. A global operation is performed on the data of each task and the result is delivered to a single target task. Operations like sum, minimum or maximum are available. Beyond that user-defined operations are supported.

MPI_Alltoall. This routine sends the data of each task to all other tasks. The routines of the *All-to-All* class are usually the most expensive ones in terms of performance and should be avoided if possible.

One-Sided Communication. The point-to-point and the collective communication are powerful tools to implement a vast variety of communication patterns. However, there might be situations where these communication modes impose drawbacks on the performance of communication in certain algorithms, for example, in cases where the sender needs to send data to another task but does not know to which and/or which data needs to be send. Or the receiver does not know that he should receive data. Another example is the case where the receiver would need to initiate the data-exchange process. All these scenarios cannot be realized with point-to-point or collective communication in a straightforward manner.
Here, the one-sided communication, also called *remote memory access* (RMA) offers a solution. In this type of communication a task (called *origin* in MPI) can access remotely the memory of one or more other tasks (*target*) and pull or push the data from or to the target. The target does not need to issue a corresponding send or receive call. However, while in explicit message passing an implicit synchronization takes place, in one-sided communication explicit synchronization is needed to indicate the completion of the data exchange. In MPI RMA communication this is realized via the *window* concept. A window is a block of memory of a task opened for remote access through MPI RMA operations.
The RMA operations are realized with calls to **MPI_Get**, **MPI_Put**, and **MPI_Accumulate** routines, with which tasks pull or push data.
Finally, MPI offers two different ways for synchronization in one-sided communication: (i) *Active Target Communication* (ATC) and (ii) *Passive Target Communication* (PTC). In ATC collective **MPI_Win_fence** calls are used which are executed on the origin as well as on the target, while in PTC **MPI_Win_lock** and **MPI_Win_unlock** routines are use only on the origin. Further routines like for *general active target communication* (GATC) are also available.

3.4 MPI I/O

Running on several thousand or more cores the I/O can become a challenging task. Input data needs to be read in to all of the tasks and the results need to be stored on disk. The amount of data to be read in or to be stored can become in simulations on nowadays HPC systems several hundred Terabytes. In principle there are three ways to implement the I/O

1. **Serial I/O**. Only one task performs all I/O operations. This model is easy to implement, however, it has some serious performance drawbacks. The data is usually distributed over all tasks. Therefore, the data has to be gathered on one task first. This introduces additional communication. Furthermore, the I/O bandwidth is limited to one task and can become a serious bottleneck. Finally, while one task performs the I/O all other tasks might idle, because they might need to wait for this task to complete reading in and distributing the data.

2. **Task-local I/O**. In this model each task has its own file for performing I/O operations. The advantages are that it is very easy to implement, the I/O needs not to be coordinated between the tasks like in the previous model and no false sharing of file system blocks occurs. However, very serious performance problems practically forbid the usage of this I/O model on HPC systems: (i) the number of files becomes quickly unmanageable when using several thousand tasks, (ii) files need to be merged to create a canonical dataset, and (iii) the file system might serialize meta data modifications. Especially due to (iii) this model is extremely expensive on parallel HPC systems and should no be used.

3. **Parallel I/O**. To overcome the drawbacks of the serial and task-local I/O, several or all tasks could write to the same file. This way the bandwidth is increased since the I/O is not restricted to one task and only one file is needed. However, one needs to carefully avoid other performance issues. For example, the access to the file must be coordinated among the tasks to avoid time penalties and eventually serialization of the I/O and the file layout might induce false sharing of file system blocks.

MPI offers routines to efficiently implement *parallel I/O* in parallel programs. Conceptually, writing to or reading from a file is treated the same way as sending or receiving messages. Files are opened and closed using the collective routines **MPI_File_open** and **MPI_File_close**. In order to provide MPI hints about the I/O to be performed and/or the file system and therefore, to optimize the I/O, **MPI_Info** objects can be used. The basic concept for the access of the file is the *view*. A view defines the data elements of a file visible to a task . Each task has an individual view of a file, which can be altered during the runtime of an application to enable different access patterns. MPI offers various routines to write to or read from a file including blocking and nonblocking, collective and noncollective routines. Three types of file access are provided:

1. **Individual file pointers**. Each task has its own file pointer that is only altered on access of that specific task.

2. **Shared file pointers**. This file pointer is shared among all tasks in the communicator used to open the file. It is modified by any shared file pointer access of any task.

3. **Explicit offset**. No file pointer is used or modified. An explicit offset is given to determine access position.

In order to optimize the parallel I/O, as many information as possible should be given to MPI. For example, collective I/O should be used instead of individual writes or reads, data should be transferred in large chunks instead of small ones and *MPI_Info* objects should be used to pass hints to MPI. This holds not only for the I/O but also for performing communication in general. This introduction to MPI can cover only very briefly the basic principles and routines. For further information the user is referred to the MPI standard [2] or to the books of Gropp, Lusk and Skjellum [9, 10]. A good introduction to programming with MPI (and OpenMP) in C is given by M. Quinn [7].

4 Open Multi-Processing (OpenMP)

While MPI is extensively used to enable parallelization of distributed memory systems, OpenMP has become extremely popular when parallelization needs to be realized on shared memory architectures. Most modern supercomputers show a hybrid distributed-shared memory architecture, i.e. MPI as well as OpenMP can be combined to get the most out of this kind of systems. OpenMP stands for *Open Multi-Processing* and offers its functionality via pragmas recognized by the compiler, runtime library routines and environment variables. In contrast to MPI, it is not necessary to link against certain libraries but OpenMP is implemented in the compiler itself and can be activated by a compiler specific option. OpenMP is based on threads. A thread is a lightweight version of a process with its own memory for storing dynamic and local data

and access to a shared (global) address space. Fig. 8 illustrates the fork-join model used by OpenMP.

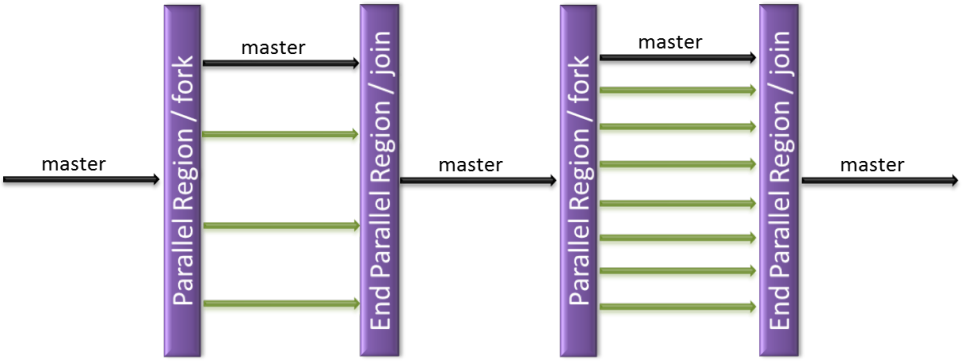


Fig. 8: *OpenMP supports the fork-join programming model. At the start of the application the master thread executes sequentially the program. When it comes upon a parallel region a team of threads is being created and joined at the end of that region.*

The program starts as usual with a single thread called the *master thread* which sequentially goes through the instructions. Once the thread meets an OpenMP pragma a team of threads is forked which execute those parts of the code that follow after this directive. The threads can act in several ways, whereby their behavior is further specified by features OpenMP provides. At the end of a parallel region which is indicated by a specific directive in Fortran code - in C/C++ the parallel region is included in curled brackets - the threads will be destroyed and the master thread proceeds with the sequential execution of the code. The following subsections will explain in more detail how OpenMP can be used to parallelize code blocks the way the programmer expect it to do. From the described strategy about how OpenMP generally works it becomes clear that it also follows the SPMD principle, i.e. one single program is started on a single node spawning several threads when needed. These threads are always numbered from 0 up to $n-1$ with n equal to the number of threads and can therefore unambiguously be identified. All these threads "see" the same program and the programmer decides what has to be done by which thread. OpenMP supports C/C++ and Fortran. While C and C++ use exactly the same constructs they are slightly different in Fortran. Listing 1 and 2 show a simple *Hello World* example to show the general usage of OpenMP and in addition to that the different naming of the directives in C/C++ and in Fortran.

Listing 1: *C example: Hello World*

```
1 #include <stdio.h>
2
3 int main() {
4 #pragma omp parallel
5   {
6     printf ("Hello World\n");
7   }
8   return 0;
9 }
```

Listing 2: *Fortran example: Hello World*

```
1 PROGRAM HELLO
2
3 !$OMP PARALLEL
4   write(*,*) "Hello World"
5 !$OMP END PARALLEL
6 END PROGRAM HELLO
```

The program is sequentially executed by the master thread until it comes upon the **omp parallel**

directive at line 4 in Listing 1 and line 3 in Listing 2 respectively. Here a predefined number of threads will be created where all of them enter the parallel region in order to execute the code. Hints on how to influence the number of threads to be used will be explained a bit later. In C/C++ the parallel region is included in curled brackets while in Fortran the directive **omp end parallel** indicates the end of the region. Here, all but the master thread are destroyed.

4.1 OpenMP Directives

In the following subsections the most commonly used OpenMP constructs are described - generally showing first the C/C++ bindings followed by the Fortran bindings - to give the reader an insight into the usage of OpenMP. For more detailed information, the OpenMP standard [3] is a suitable choice of source.

#pragma omp parallel, !$omp parallel is the integral construct which starts parallel execution. Here, a team of threads is spawned to execute the region. The team remains active until the end of the parallel region is reached by all of them ensured by an implicit barrier. Typically further OpenMP constructs are included in this region to steer the behavior of the threads in dependence of the work they are expected to deal with.

The available constructs can be divided into different classes in dependence of the behavior they imply. This introduction gives an overview about *work sharing* constructs and *master and synchronization* constructs. The most frequently used *work sharing* constructs are:

#pragma omp for, !$omp do The iterations of the loop are distributed over the available threads. The way how the iterations are divided among the threads can be specified by the **schedule** clause. Oftentimes the workload of each iteration is more or less the same so that a static schedule might give the best performance as each thread gets the same number of iterations. If the work load differs from iteration to iteration then further *schedule clause kind* values can be chosen to meet the demands. When loops are parallelized with OpenMP the calculations inside an iteration must not depend on other iterations, this might lead to wrong results. In order to achieve best performance of loops with OpenMP, some restrictions apply to simplify compiler-based parallelization. One of the rules impose the restriction that the form of the loop allows to compute the number of iterations prior to entry into the loop, from which one can deduce that it is not possible to parallelize *while* loops. Besides that the program must complete all iterations of the loop, i.e. in C/C++ it is not allowed to use *break* or *goto* in Fortran code this means avoiding *exit* and *goto*. However, it is allowed to exit the current iteration and go over to the next one with *continue* in C/C++ and *cycle* in Fortran. Finally, the termination of the entire program inside a loop is possible. All in all a *for/do* loop can only have a very restricted form as shown in Listing 3.

Listing 3: *Parallel loop for C/C++. The same restrictions hold true for Fortran*

```
#pragma omp for (var=first;var cmp_op end;incr_expr)
 for (...) {loop-body}

cmp_op: <, <=, >, >=
```

```
incr_expr: ++var, var++ , --var, var--
           var += incr, var -= incr
           var = var + incr, var = incr + var
```

#pragma omp sections, !$omp sections. Within *sections* sub-blocks called *section* can be defined whereby one thread per sub-block executes the included code. If there are more blocks than threads, threads will be assigned to new blocks after finishing their work in the current blocks.

#pragma omp single, !$omp single. This directive allows to execute some code in a parallel region by only one thread. An implicit barrier at the end of the *single* region ensures that the remaining threads do not proceed until this region is finished.

After mentioning some of the general work sharing constructs, now a description of some master and synchronization constructs follows:

#ragma omp master, !$omp master are similar to the *single* directive but there are two differences, namely the region is exclusively executed by the master thread and in addition to that no implicit barrier is set at the end of this region.

#pragma omp critical, !$omp critical are executed by all threads but only one thread after the other. The order of the threads executing the critical region is undetermined. Critical regions can be named so that all critical regions with the same name build a group. Now, if a critical region is entered by a thread x, further threads in front of critical regions with the same name will wait until thread x leaves its critical region. If no name is given all threads next to critical regions will idle as long as a thread is active in any critical region. In many cases this behavior is unintended, as data in different critical regions do not necessarily depend on each other and therefore causing no problems when executed concurrently.

#pragma omp barrier, !$omp barrier threads are only allowed to proceed when all prior tasks bound to the barrier are finished.

#pragma omp atomic, !$omp atomic ensures that a specific memory location is updated atomically. In dependence of the OpenMP implementation *atomic* might be replaced with a *critical* construct. However, the *atomic* construct permits better optimization as it is based on hardware instructions.

In addition to directives OpenMP also offers so called *combined* directives for parallel regions which only contain a single parallel loop or parallel sections. These directives are useful shortcuts but they do not bring further functionalities. Listings 4 and 5 illustrate their usage.

Listing 4: *C/C++: Combined constructs for parallel for loops and parallel sections*

```
  #pragma omp parallel for \
      [ParameterList]
    for-loop

#pragma omp parallel sections\
      [ParameterList]
  {
    [\#pragma omp section]
      code block
    [\#pragma omp section]
      code block
    ...
  }
```

Listing 5: *Fortran: Combined constructs for do loops and parallel sections*

```
!$omp parallel do &
!$omp [ParameterList]
  do loop
[!$omp end parallel do]

!$omp parallel sections &
!$omp [ParameterList]
[!$omp section]
  code block
[!$omp section]
  code block
...
!$omp end parallel sections
```

There is a whole bunch of further constructs that cannot be mentioned at this point. For further reading it is recommended to consult the OpenMP standard [3]. Furthermore the books of B. Chapman et al.[11] and R. Chandra et al.[12] give a decent introduction to the subject. For those who are interested in implementing OpenMP pragmas in their C++ codes the book of S. Hoffmann and R. Lienhart [13] gives hints to this topic.

4.2 Data Sharing Attribute Clauses

The OpenMP user needs the possibility to influence the scope of variables which appear in parallel regions. In OpenMP this is enabled through so-called *clauses*, which can be added to directives. Apart from some exceptions variables are shared per default in parallel regions. This means, all threads of a team share the same memory space and have read and write access to stored data. However, because of data consistency reasons it might be necessary to assign certain variables at least temporarily to threads, i.e. from time to time threads need local copies of otherwise shared variables. The most important clauses are:

shared(*var-list*), where *var-list* is a comma-separated list of variables used in the parallel region, i.e. there is only a single instance of these variables in shared memory which are accessible by all threads.

private(*var-list*), where again *var-list* is a comma separated list of variables which are intended to be used in the parallel region. Here, each thread allocates its own private copy of the data. However, these local copies only exist during the execution of the parallel region. Furthermore, the values of these variables are undefined upon entry and exit of parallel regions.

firstprivate(*var-list*). Variables specified in the *firstprivate* clause are also declared to be private like in the *private* clause. However, in addition to that these variables get initialized with the value the original shared variable had just before entering the parallel region.

lastprivate(*var-list*) declares variables as private and the value of the last thread finishing the parallel region is written to the corresponding shared variable afterwards. Variables can appear concurrently in both clauses, *firstprivate* and *lastprivate*.

C++: default(shared|none), Fortran: default(private|firstprivate|shared|none) sets the default data mode, whereby private is only allowed in Fortran. Apart from loop index variables in parallel loops and local variables of subroutines called inside parallel regions, all values are shared by default.

4.3 Runtime Library Functions and Environment Variables

OpenMP provides plenty of helpful routines to gather information from the thread environment and in addition to that it provides functions to adapt the environment to the needs of the programmer. Subsection 4.3 presents an overview of the most important ones to give a good starting point for their usage. Before these functions can be used in C/C++ codes one needs to include the corresponding header as show in Listing 6. For Fortran codes load the corresponding module or alternatively include the OpenMP related header like in Listing 7.

Listing 6: *C/C++: Activate OpenMP Runtime Libraries for usage in the code*

```
#include <omp.h>
```

Listing 7: *Fortran: Activate OpenMP Runtime Libraries for usage in code*

```
include "omp_lib.h"
! alternatively: use omp_lib
```

In addition to that a large number of environment variables are at hand to influence the OpenMP thread environment. Where possible they are presented along with the corresponding runtime functions. These variables need to be set in the underlying shell. Generally, settings established by environment variables are predominated by settings applied via runtime routines.

omp_set_num_threads(*num-threads*) sets the number of threads to be used in subsequent parallel regions. *num-threads* must evaluate to a positive number.
OMP_NUM_THREADS. The value of the variable must be a positive number.

omp_get_num_threads() returns the number of threads in the current team. This function needs to be called within a parallel region to give the actual number of threads. Outside a parallel region only the master thread exists, therefore *omp_get_num_threads()* would give back one.

omp_get_thread_num() returns the thread's number. Like MPI tasks threads are numbered from 0 to $n-1$, with n equal to the number of threads. Thread 0 is always the master thread. The concept of *SPMD* becomes obvious as only a single instance of a program is executed but in dependence of the thread number the code can branch and different threads work concurrently on different parts.

omp_get_wtime() offers a convenient way to determine the wall time the execution of a code block consumes. *omp_get_wtime()* returns the elapsed wall clock time since a certain point of time in the past. Therefore, the difference between two calls gives the elapsed time between the calls:

Listing 8: *Measuring runtimes*

```
double start_time = omp_get_wtime();
// code block
double elapsed_time = omp_get_wtime()-start_time
```

omp_set_nested(nested), OMP_NESTED. Nesting of parallel regions are allowed, but not all compilers implement this feature. The variable *nested* must evaluate to an integer in C/C++, in Fortran codes it is a logical variable. If available this feature can be switched on and off by *OMP_NESTED*. The value of this variable must be true (in C/C++: $\neq 0$) or false (in C/C++: 0). There is another function called *omp_get_nested()* which informs whether the nested feature is enabled or not.

5 Conclusion

In this article the basic concepts and models of parallel computing were introduced. The principles of modern parallel hardware were discussed and parallel programming models and concepts were presented, including a design strategy for parallel programs. Efficiency and scalability considerations were explained and Amdahl's law was introduced as a helpful means in order to estimate the quality of the parallelization of algorithms. Finally, the basic concepts of MPI and OpenMP were introduced.
This introduction to parallel programming is of course by no means exhaustive but the covered ideas and presented features should help to make a first step towards parallel programming with MPI and/or OpenMP.
The current MPI standard contains more than 400 routines and only a few very basic routines and few features of MPI could be presented. Important concepts like groups, contexts and communicator could only be briefly touched. For others like the very important topics *derived datatypes* and *process topologies* the reader is referred to the MPI standard. In order to achieve a good parallel performance using MPI a balance between computation and communications on each task should be kept and as much information as possible should be provided to MPI, so that MPI is able to optimize the communication.
For OpenMP some further topics worth to look at after getting familiar with the general concepts are *tasks*, for instance, which give high flexibility when blocks of code need to be dynamically identified and executed in parallel. To speed up the code it might be useful in certain cases to switch off implicit barriers which are bound to some directives. Here, the *nowait* clause comes in handy. Sometimes situations might occur in codes that cannot be parallelized adequately with the general OpenMP directives, so that a detailed look at *locks* might help. *Locks* allow an exclusive access of single threads to a resource on a fine-grained level. However, special care needs to be taken by the programmer as OpenMP does not care of proper locking and unlocking of resources at this point so that there is always the potential danger of data corruption and deadlocks.
After presenting several techniques of parallel programming it might be useful to end with some general hints that help to cope with the problem to convert a serial program into a well performing parallelized application:

- General hints
  - Check that the serial code is well optimized.

  - Seek for compiler settings which might increase the performance of the code.
  - Estimate the scalability with the help of Amdahl's law.
  - Find parts of the code which consume the most computation time.

- MPI-related hints
  - Try to avoid communication.
  - Avoid many small messages, try to combine them into few larger messages.
  - Provide information to MPI (collective routines, *MPI_Info* objects).
  - Avoid deadlocks, use nonblocking communication.
  - Do not send many messages without posting corresponding receives, internal buffers might fill up and cause deadlocks.
  - Check the environment where your application runs, especially the communication protocol limits (*eager* and *rendezvous* protocol) of you MPI implementation.

- OpenMP-related hints:
  - Reduce the amount of parallel regions to a minimum.
  - Optimize the most outer part of nested loops.
  - Use the *nowait* clause whenever possible.
  - Balance the workload over all threads.
  - Avoid false sharing effects.
  - Name all critical regions.
  - Consider the environment in which the application runs.

6 Appendix

Last but not least the appendix provides an example code illustrating some of the described concepts to increase the performance by parallelization. Eq. (6) - (8) show a way to compute the value of π by numerical integration.

$$\pi = 4.0 \int_0^1 \frac{1.0}{1.0 + x^2} dx \approx \sum_{i=1}^{n} \frac{4.0}{(1.0 + x^2)} \Delta x \tag{6}$$

$$\Delta x = \frac{1}{n} \tag{7}$$

$$x = (i - 0.5)\Delta x \tag{8}$$

Besides the serial version of the code in Listing 9 a hybrid one using MPI functions and OpenMP directives as well is given in Listing 10.

Listing 9: *Serial C code implementing a numerical calculation of π based on Eq. (6) - (8).*

```
1  #include <stdio.h>
2  #include <math.h>
3
4  int main(int argc, char** argv) {
5
6    double ref_pi = 3.141592653589793238;
7    long iterations = 4*200000000;
8    double dx = 1.0/iterations;
9    double sum = 0.0;
10   double x, fx, pi;
11   long i;
12
13   for (i=1 ; i<=iterations ; ++i) {
14     x = dx * (i-0.5);
15     fx = 4.0/(1.0+x*x);
16     sum += fx;
17   }
18
19   pi = dx*sum;
20
21   printf("Result pi    : %20.18f\n", pi);
22   printf("Ref. val. pi: %20.18f\n", ref_pi);
23   printf("Diff. to pi : %20.18f\n", fabs(pi-ref_pi));
24
25   return 0;
26 }
```

According to Foster's design strategy for parallel programs from chapter 2.3 the four steps could be applied Listing 9 in the following way:

- *Partitioning*: line 13 - 16 , a primitive task would correspond to a single iteration of the loop.
- *Communication*: line 19, the primitive tasks are independent from each other, communication is only necessary at the very end when π is calculated.
- *Agglomeration*: several iterations of the *for* loop will be agglomerated into one task.
- *Mapping*: depending of the number of MPI processes several iterations will be performed by one MPI task and its corresponding OpenMP threads on each core.

Listing 10: *Hybrid C code implementing the computation of π based on Equations (6) - (8). Here, MPI as well as OpenMP were used for parallelization.*

```
1  #include <stdio.h>
2  #include <omp.h>
3  #include <math.h>
4  #include <mpi.h>
```

```
 5
 6 int main(int argc, char** argv) {
 7
 8   double ref_pi=3.141592653589793238;
 9   long iterations = 4*200000000;
10   double dx = 1.0/iterations;
11   double sum = 0.0, local_sum = 0.0;
12   double x, fx, total_pi, local_pi;
13   long i;
14   int ierror, nranks, myrank;
15   double start_time;
16
17
18   // Initialize MPI
19   MPI_Init(&argc, &argv);
20   MPI_Comm_size(MPI_COMM_WORLD, &nranks);
21   MPI_Comm_rank(MPI_COMM_WORLD, &myrank);
22
23   if (myrank == 0) start_time = MPI_Wtime();
24
25 #pragma omp parallel default(none) \
26 shared(myrank,iterations,nranks,dx,sum) \
27   private(i,x,fx) firstprivate(local_sum)
28   {
29 #pragma omp for
30     for (i=myrank+1 ; i<=iterations ; i=i+nranks) {
31       x = dx * (i-0.5);
32       fx = 4.0/(1.0+x*x);
33       local_sum += fx;
34     }
35 #pragma omp critical
36     {
37       sum = sum + local_sum;
38     }
39   }
40
41   local_pi = dx*sum;
42
43 MPI_Reduce(&local_pi, &total_pi, 1, MPI_DOUBLE,\
44   MPI_SUM, 0, MPI_COMM_WORLD);
45
46 if ( myrank == 0 ) {
47     printf("Result pi    : %20.18f\n", total_pi);
48     printf("Ref. val. pi: %20.18f\n", ref_pi);
49     printf("Diff. to pi : %20.18f\n", fabs(total_pi-ref_pi));
50     printf("Time needed : %7.4f sec\n", MPI_Wtime()-start_time);
51   }
```

```
52
53    MPI_Finalize();
54    return 0;
55 }
```

Both versions are brought face to face to underline the applied modifications. The *for* loop is the only CPU intensive part in the code and therefore the only part worth to parallelize. The general approach starts with including MPI routines followed by implementing the OpenMP directives. Lines 19 - 21 of Listing 10 arrange the initialization of the MPI tasks, return the number of tasks and give back the ranks for each of the tasks. Then each rank computes its part of π in the *for* loop. Finally all individual results of the tasks are combined in the variable *total_pi* by calling the collective communication function *MPI_Reduce()* in line 43. Line 46 ensures that only the master rank writes out the results.
In a second step further performance can be obtained by the use of OpenMP as parallelizing loops is one of its specialties. The parallel region for the threads is opened in line 25. Here, the default is set to *none* in order to explicitely decide which variable gets which scope (*private*, *shared* etc.). As each MPI task computes its part of π and in turn all threads that belong to one MPI task compute the sum for π, it is necessary to introduce a further variable *local_sum* to avoid data dependencies. *local_sum* is 0 on entry of the parallel region and each thread calculates its sum. Afterwards all these partial sums of all threads need to be summed as shown in line 37. Doing this in parallel involves the danger of wrong results as several threads might read and write *sum* at the same time mutually overwriting their results. Therefore, this summation is processed in a critical region to assure that only one thread adds its value at a time.

References

[1] http://www.top500.org/

[2] The MPI Forum: The Message Passing Interface, http://www.mpi-forum.org/

[3] OpenMP Application Program Interface, Version 4.0 - July 2013 http://www.openmp.org/mp-documents/OpenMP4.0.0.pdf.

[4] J. von-Neumann (1945), http://systemcomputing.org/turing%20award/Maurice_1967/TheFirstDraft.pdf.

[5] M. Flynn: Some Computer Organizations and Their Effectiveness, IEEE Trans. Comput., Band C-21, S. 948960, 1972.

[6] I. Foster: Designing and Building Parallel Programs: Concepts and Tools for Parallel Software Engineering. Reading, MA: Addison-Wesley, 1995.

[7] M. J. Quinn: Parallel Programming in C with MPI and OpenMP. New York, NY: McGraw-Hill, 2003.

[8] G. Amdahl: Validity of the single processor approach to achieving large scale computing capabilities. In: AFIPS Conference Proceedings, 30, 483-485, Washington, D.C.: Thompson Books, 1967.

[9] W. Gropp, E. Lusk, A. Skjellum: Using MPI: Portable Parallel Programming with the Message Passing Interface, 2nd edition. Cambridge, MA: The MIT Press, 1999.

[10] W. Gropp, E. Lusk, A. Skjellum: Using MPI-2: Advanced Features of the Message Passing Interface. Cambridge, MA: The MIT Press, 1999.

[11] B. Chapman, G. Jost, R. van der Pas: *Using OpenMP*, MIT Press, 2008

[12] R. Chandra, L. Dagum, D. Kohr, D. Maydan, J. McDonald, R. Menon: *Parallel Programming in OpenMP*, ACADEMIC PRESS, 2001

[13] S. Hoffmann, R. Lienhart: *OpenMP*, Springer-Verlag Berlin Heidelberg, 2008

D 3 Numerical methods for the eigenvalue problem in electronic structure computations [1]

Edoardo Di Napoli
Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercomputing" (Forschungszentrum Jülich, 2014).

1 Introduction

Every Density Functional Theory (DFT) method is grounded on a variational principle directly inspired by the fundamental theorem of Kohn and Hohenberg [1], and its practical realization [2]. Central to DFT is the solution of a large number of coupled one-particle Schrödinger-like equations known as Kohn-Sham (KS)

$$\left(\frac{\hbar^2}{2m}\nabla^2 + \mathcal{V}_{\text{eff}}[n(\mathbf{r})]\right)\phi_i(\mathbf{r}) = E_i\phi_i(\mathbf{r}) \quad ; \quad n(\mathbf{r}) = \sum_i f_i\phi_i(\mathbf{r}).$$

Due to the dependence of the effective potential $\mathcal{V}_{\text{eff}}$ on the charge density $n(\mathbf{r})$, in itself a function of the orbital wave functions $\phi_i(\mathbf{r})$, the KS equations are non-linear and are generally solved by a sequence of self-consistent field (SCF) cycles.
The KS equations need to be "discretized" in order to be solved numerically. Intended in its broadest numerical sense, the discretization translates the KS equations in a non-linear eigenvalue problem. Eigenproblems generated by distinct discretization schemes have numerical properties that are often substantially different; for sake of simplicity we can group most of the schemes in three classes. The first and the second classes expand each of the one-particle orbital wave functions $\phi_i(\mathbf{r})$ appearing in the KS equations on a specific set of basis functions

$$\phi_i(\mathbf{r}) \longrightarrow \phi_{\mathbf{h},i}(\mathbf{r}) = \sum_{\mathbf{G}} c^{\mathbf{G}}_{\mathbf{h},i}\psi_{\mathbf{G}}(\mathbf{h},\mathbf{r}), \tag{1}$$

where $\mathbf{G}$ and $\mathbf{h}$ represent cumulative general sets of indices.
The first class makes use of simple plane waves $\psi_{\mathbf{G}}(\mathbf{k},\mathbf{r}) \sim e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}$ with **G** and **k** (here the more commonly used **k** replaced **h** index) being vectors in momentum space. This basis set is very simple to handle in Fourier space where the kinetic energy operator $\frac{\hbar^2}{2m}\nabla^2$ is diagonal. On the contrary, the potential term $\mathcal{V}_{\text{eff}}[n(\mathbf{r})]$ gives rise to a large number of off-diagonal terms. Moreover, when close to the origin of the atomic Coulomb potential, plane waves can oscillate quite wildly giving rise to computational difficulties. For the latter reason plane waves are usually utilized in combination with pseudo-potentials where the singular part of the Coulomb term is replaced with a softer function emulating the screening effects of the core electrons. The resulting pseudo-potential contains both local and non-local terms inducing dense eigenproblems.
The second class resorts to localized functions $\psi_{\mathbf{G}}(\mathbf{h},\mathbf{r}) \sim R_\ell(\mathbf{r}_a)Y_{\ell,m}(\hat{\mathbf{r}}_a)$ which combine radial functions around an atom with spherical harmonics. One popular example are the Gaussian type orbitals (GTO)[3, 4]. In this basis set the radial functions are equal to $R_\ell(\mathbf{r}_a) = r_a^\ell e^{-\alpha_p r^2}$, where $\mathbf{G} = (a,\ p)$ is indexing the atom-localized primitive Gaussians, **h** parametrizes quantum numbers ℓ and m, and $c^{\mathbf{G}}_{\ell,m,i}$ include the contraction coefficients and normalization constants. The GTOs set is particularly convenient since it capitalizes on the fact that integrals of multiple products of Gaussians can be easily reduced, in a chain of simplifications, to products of single Gaussian integrals. On the opposite many functions are needed to represent faithfully each electronic orbital, ending in an expensive bookkeeping process. The net result is that methods based on GTOs, as well as other kind of localized orbitals, are usually quite accurate but require a big deal of optimization to ensure the basis set is complete.
A special place is occupied by methods based on Linearized Augmented Plane Waves (LAPW) [5, 6], where a mix of radial functions and plane waves are used. The main advantage of these methods reside in the ability of employing the full potential without the need of distinguishing

between core and valence electrons. For this reason this basis set is usually considered quite accurate – even if expensive – for the simulation of transition metals. Since they give rise to dense problems, for the purpose of our classification, we just include LAPW-based methods in the first class.
Methods in the third class do not use an explicit basis set but discretize the KS equations using functions centered on a uniform mesh in real space. The easiest real-space methods use high-order finite differences [7] but there are also implementation using finite elements or wavelets [8, 9]. In the latter case, it is possible to have an adaptive mesh closer to the nucleus using the scaling properties of wavelets. These methods use functions that are localized, a technique which allows the development of order n methods. In this formalism the potential entries in the Hamiltonian matrix decay quite radiply away from the main diagonal while the size of the matrix is proportional with the total number of grid points leading to quite large eigenproblems.
Eigenvalue problems emerging from the first two discretization classes consist of **dense** matrices of **small-to-moderate** size while, within real space methods, one ends up with **very large** and **sparse** matrices. While for most DFT methods only a **fraction** of the eigenspectrum is required, the magnitude of such fraction can vary wildly from method to method (even within the same class). In addition, depending on the choice of basis set, the eigenproblems could be either standard or generalized. In the latter case the numerical properties of the overlap matrix strongly depend on the over-completeness of the basis set. In some cases this amounts to deal with eigenproblems with large **condition** number.
Due to the dramatically different set of properties of the eigenproblems, each DFT method uses a distinct strategy in solving for the required eigenpairs. For instance it is quite common that methods ending up with dense problems and a fraction of the eigenspectrum larger than 1% use direct eigensolvers. On the opposite many real space methods end up with very sparse and quite large matrices which are not even generated explicitly. Consequently these methods make use of iterative eigensolver based on Krylov- or Davidson-like subspace constructions. From the point of view of software packages for distributed memory architectures, the choice between direct or iterative eigensolvers leads to the use of traditional parallel libraries like, respectively, ScaLAPACK [10] or PARPACK [11]. Not satisfied with traditional libraries, some of the newer implementations ended up developing their own tailored eigensolver (e.g. [12, 13]).
In the following section we will address both direct and iterative eigensolvers, their properties, general parallel implementation strategies as well as some tailored algorithms. Due to the vastness of the subject we will not be exhaustive but will use specific algorithmic examples to illustrate the differences in their computational approach. The careful reader can expect to gain some general insight on which algorithm is best suited to which DFT implementation.

2 Definitions and tools

Let us first introduce some definitions and terminology which are common among specialists and will be used in the rest of the chapter. We define three distinct type of eigenproblems. They are all identified by the equation

$$Ax = \lambda Bx \tag{2}$$

where $A,\ B \in \mathbb{C}^{n\times n}$ are given matrices[2] and one seeks the unknown scalars $\lambda \in \mathbb{R}$ and the associated vectors $x \in \mathbb{C}^n$. The latter are referred to as, respectively, *eigenvalues* and

[2]In this chapter we always use the letter n to indicate the size of the matrices

eigenvectors and are usually displayed as an *eigenpair* $(\lambda,\ x)$. It is also common practice to designate the pair $(A,\ B)$ as an *eigenpencil*.
In the most general case both $A,\ B$ are Hermitian symmetric ($A^\dagger = A,\ B^\dagger = B$) and B is positive or negative definite[3]. In this case we will refer to Eq. (2) as the *generalized Hermitian eigenvalue problem*, GHEVP in brief . If $B = I$ then Eq. (2) reduces to the *standard Hermitian eigenvalue problem* (HEVP) $Ax = \lambda x$. In the particular case where A is also tridiagonal ($a_{ij} = 0\ \ j > i+1$) the eigenproblem is identified as *symmetric tridiagonal eigenvalue problem* (STEVP). In the latter case all non-zero entries of the matrix A are real-valued.
It is important that B is a definite matrix for the GHEVP to be well-posed. As we already mentioned in the Introduction a measure of such property is the *condition number*. This number is defined as

$$\kappa(B) \doteq \frac{\|B\|_2}{\|\ B^{-1}\ \|_2} = \frac{\max_i(\sigma_i)}{\min_j(\sigma_j)} = \frac{|\lambda_{\max}(B)|}{|\lambda_{\min}(B)|}.$$

It can be expressed in relation to the singular value decomposition of $B = W\Sigma V$ where $W,\ V$ are two unitary matrices and $\Sigma = \mathrm{diag}(\sigma_1 \dots \sigma_n)$ is a diagonal matrix holding the singular values of B. In the case of GHEVP such values corresponds just to the absolute value of the largest and smallest eigenvalues of B. Consequently a large condition number implies that B has one or more eigenvalues very close to zero. When this happens some of the eigenvectors of the GHEVP are close to linearly dependent and B can be hardly inverted or factorized. Consequently it becomes very difficult to solve the eigenproblem.

Metrics and performance — The general concept of performance can be associated with several metrics. One could consider the total *CPU time* to completion as the measure of the performance of an algorithm implementing an algebraic operation. In many cases this is considered an important performance metric to compare two algorithmic variants of the same operation tested in exactly the same condition.
Alternatively one could instead decide that the best algorithm is the one performing the least number of elementary operations. One refers to the number of elementary operations of a algebraic transformation as its *complexity*. For example the complexity of a real-valued matrix-vector product is $2n^2-n$ (n multiplications and $n-1$ additions per vector entry). It is customary to consider only the leading term in n contributing to the complexity. For the matrix-vector product the complexity would then be indicated as $\mathcal{O}(n^2)$. The complexity of a matrix operation is directly correlated with the number of *floating point operations* (flop) a computer is capable to handle.
The number of flop is a more fundamental quantity with respect to the complexity of an algorithm and it is at the base of the modern way of judging the efficiency of a numerical computation. This is due to the fact that some elementary operations costs more (in terms of flop counts) than others. In this chapter, as it is done in current literature, we adopt as definition of *performance* the flop count rate which is defined as

$$\text{PERFORMANCE}\ (Flops) \doteq \frac{\text{NUMBER OF OPERATIONS}}{\text{TIME}\ (sec)}.$$

[3]It is common practice in numerical linear algebra to represent the hermitian conjugation with the letter H (or T in the case of real symmetric matrices). Contrary to this habit we use in this chapter the symbol †, commonly used among physicists.

Since it is important to put in perspective the absolute performance with the computing architecture used, we also introduce the *theoretical peak* performance of a machine defined as

$$\begin{array}{c}\text{THEORETICAL}\\ \text{PEAK PERFORMANCE}\end{array} \doteq (\text{FREQ.}) \cdot (\#\ \text{OF CORES}) \cdot \frac{\#\ \text{OF ELEM. OPERATIONS}}{\text{CYCLE}}.$$

With this definition in hand a better measure and a useful tool is the *efficiency* η of a routine, defined as its performance over the theoretical peak performance. For example one may choose an expensive (in terms of flop counts) but efficient algorithmic variant respect to a cheaper but less efficient one. If the efficiency of the former can compensate for the higher complexity, the first routine will be faster in terms of CPU time. Clearly this is not the only consideration that should guide the final user in choosing an algorithm.
Another important metric to consider is the amount of memory necessary to perform a specific computation. The *memory requirements* of an algorithm – also known as *workspace* – is the amount of main cache memory which is needed to complete the computations of a certain algorithm. For example one may want to choose a more "green" computational approach and minimize the movement of data between memory and processors[4] at the cost of some performance. In fact memory has became an expensive commodity both in terms of memory size available per core and consumption of energy. Moreover some algorithms may require a large chunk of cache memory reserved for workspace: if a specific architecture have limited memory per core such algorithms may be bound by the size of the input data they can handle.
Last but not the list we want to mention *accuracy* among the possible metrics which are used to analyze the performance of an algorithm implementation. Some algorithms may, in fact, compromise on accuracy in order to improve the performance. Moreover some algorithms are intrinsically more stable and accurate than others; a property that can play an important role in choosing the correct algorithm for a specific application. When dealing with eigensolvers, accuracy of the results is measured by the value of the *residual norm* and *numerical orthogonality* defined, for the HEVP, as

$$r(\lambda, x) \doteq \max_i \frac{\parallel Ax_i - \lambda x_i \parallel}{\parallel A \parallel n\epsilon} \qquad ; \qquad O(x) \doteq \max_{i \neq j} \frac{\|x_j^\dagger x_i\|}{n\epsilon}.$$

Here ϵ indicates the relative machine precision. The above definition can be generalized in an obvious manner to all the other eigenproblems.

Parallelism and scalability — When a program is run on more than one processor, the metrics above are not enough to describe the "quality" of an algorithm and its implementation in a routine. What is lacking are tools addressing the ability of the algorithm to run efficiently in a parallel fashion. The *speed-up* of a routine provides a first tool in this direction. Speed-up can be defined in several ways and we refer to the works of Amdahl and Gustavsson for a theoretical oriented discussion [14, 15]. For our practical purposes we define below the speed-up for *strong scalability* and *weak scalability*

$$\text{SPEED-UP}_{strong} = \zeta_s \doteq \frac{t_{ref}(n)}{t_p(n)} \qquad ; \qquad \text{SPEED-UP}_{weak} = \zeta_w \doteq \frac{t_{ref}(n)}{t_{\beta p}(\alpha(n))}. \tag{3}$$

In the above equations, $t_{ref}(n)$ and $t_p(n)$ indicate the execution time measured in seconds for a reference hardware (e.g. one core) and p processors respectively. Strong scalability measures

[4]Data movement is by far the most energy expensive process in a computation

the performance of an algorithm when the number of processors is increased keeping fixed the size of the data (so the flop count). A perfectly scalable algorithm then should have its execution halved each time the number of processors in doubled.
Almost every routine contains portions that are intrinsically serial and consequently will never scale. For this reason a better measure of the effective scalability of an algorithm is its ability to process larger sets of data (in our case eigenproblems of increasing size) as more processors are available. Weak scalability parametrizes this property by showing the speed-up an algorithm achieves when increasing the size of the system and proportionally augmenting the number of processors. What is kept constant is the flop count per processor. As a consequence the constant β is a function of α. For example if the complexity of an algorithm is $\mathcal{O}(n^3)$ every time n is doubled the number of processor has to be increased by a factor of $\beta = 2^3$.
Similarly to the performance efficiency η, we can introduce the *parallel efficiency* for both strong and weak scalability

$$e_s \doteq \frac{t_{ref} \cdot p_{ref}}{t_p \cdot p} \qquad ; \qquad e_w \doteq \frac{t_{ref} \cdot p_{ref} \cdot \alpha_{ref}(n)}{t_p \cdot p \cdot \alpha_p(n)}$$

As for η these definitions help normalizing the scalability of the routine under scrutiny. In other words perfect scalability corresponds to an horizontal line in correspondence of the dimensionless value 1.
Two important concepts for codes which run on parallel architectures are the *algorithmic block sizes* and the *distribution block sizes*. In the first case one refers to the sizes $m_b \times n_b$ of the blocks $A_{s,t}$ a matrix A of arbitrary data type is partitioned in

$$A = \begin{pmatrix} A_{0,0} & \dots & A_{0,N-1} \\ \vdots & & \vdots \\ A_{M-1,0} & \dots & A_{M-1,N-1} \end{pmatrix},$$

with the exception of boundaries blocks which can be smaller. Similarly the total number of p processes involved in the computation are logically viewed as a two-dimensional cartesian grid having distribution block sizes r and c with $p = r \times c$. Each block $A_{s,t}$ is is distributed over the grid in such a way that the process (s, t) owns, in a contiguous manner, the blocks

$$A = \begin{pmatrix} A_{\gamma,\delta} & A_{\gamma,\delta+c} & \dots \\ A_{\gamma+r,\delta} & A_{\gamma+r,\delta+c} & \dots \\ \vdots & \vdots & \end{pmatrix},$$

where $\gamma \equiv (s + \sigma_r) \mod r$ and $\delta \equiv (t + \sigma_c) \mod c$, and σ_r and σ_c are arbitrarily chosen alignment parameters. We will see in later sections how both these concepts are crucial for high-performance computing.

3 Direct eigensolvers

Eigensolvers are categorized by the choice of approach that goes from the input matrices A and B, defining the eigenproblem, to the eigenpairs $(\lambda,\ x)$ characterizing its solution. For direct eigensolvers this path goes through the *direct* diagonalization of the eigenproblem matrices.

In other words, as part of the solution process, each matrix defining the eigenproblem is algebraically manipulated so as to bring it to diagonal form. In such form its diagonal elements correspond to the eigenvalues of the problem whose eigevectors can be subsequently computed. The diagonalization process is usually carried on with a series of transformations which maintain the symmetry properties of the matrix and, most importantly, its spectrum. In the case of Hermitian eigenvalue problems this target is achieved with a series of similarity transformations which generically modify the values of every matrix entry. In particular if the matrix A, defining the problem $Ax = \lambda x$, has a substantial number of zero entries, the transformed matrix $\tilde{A} = SAS^\dagger$ is usually densely populated. Consequently the diagonalization process is convenient for eigenproblems with matrices that are already dense[5] while it is disadvantageous for sparse matrices. For the latter a method preserving the sparsity structure, and so limiting the total number of floating point operations to solution, is preferred.
Since direct eigensolvers act on all the entries of the eigenproblem matrices, the number of elementary operations, which are performed during the diagonalization, is directly proportional to the size n of the problem. We will see in the next sections that complexity, performance and memory workspace constraints are the most important parameters guiding the computational scientist in selecting the appropriate algorithm for its needs.

3.1 The stages of a direct eigensolver and their algorithmic variants

In Sec. 2 we have defined three distinct type of eigenproblems based on the properties of the matrices associated with them. These eigenproblems can also be seen as a chain of nested problems – GHEVP $\rightarrow$ HEVP $\rightarrow$ STEVP – where each type is connected with the previous one by a non-singular linear transformation which preserve the spectrum. Such linear transformation can be seen as the action of a pair of invertible matrices K and M on the eigenpencil matrices $(A, B) \rightarrow (KAM, KBM)$. In order to preserve the symmetry of the eigenpencil the transformation KXM needs to satisfy the additional requirement $M = K^\dagger$ which restrict us to similarity transformations.
In practice the path that goes from a GHEVP to the computation of its complete (or partial) set of eigenpairs can be schematically divided in six stages. Along the road we will recover the HEVP and the STEVP and their solutions, so that there is no need to describe these other two type of problems and their path to solution. Since we are dealing with direct solvers in each stage we will operate just on the input matrices (A, B) defining the GHEVP.

(i) This stage consists in factoring the B matrix in its Cholesky components $B = LL^\dagger$, where L is a lower triangular matrix. This factorization is unique and possible only if the matrix B is positive definite. A measure of positive-definiteness is given by the condition number $\kappa(B)$. If this has a very large value it may signal that $\lambda_{\min}(B) \approx 0$ making really hard to numerically compute the factorization. In some DFT methods this condition has to be verified in advance before attempting the factorization. This stage is referred to as *Cholesky decomposition*.

(ii) In the second stage the matrix L is used to perform the first linear transformation bringing B to diagonal form $B \rightarrow L^{-1}LL^\dagger(L^\dagger)^{-1} = I$ and $A \rightarrow L^{-1}A(L^\dagger)^{-1} = C$. Consequently the eigenpencil is now reduced to (C, I) corresponding to the standard eigenvalue problem $Cy = \lambda y$. While the eigenvalues are preserved by this transformation, the same

[5]The term *dense* is commonly used to address matrices having a number of non-zero entries greater than few percentage points.

cannot be said by the eigenvectors y, which are related to the original one by $y = L^{\dagger}x$ and will motivated the last stage. The second stage is known as *reduction to standard form*.

(iii) Evidently if one has to solve for just an HEVP the first two stages are redundant and one can start directly from this stage. At this point one builds a unitary transformation, parametrized by Q, which acts exclusively (the action on I is trivial) on $C \to QCQ^{\dagger} = T$ where T is a real-valued symmetric tridiagonal matrix. The eigenproblem has now been reduced to $Tz = \lambda z$. Like in the previous stage the spectrum is preserved while the eigenvectors go through another transformation $z = Q^{\dagger}y$. We will see later that there are several methods for building the unitary matrix Q. This stage goes under the name of *reduction to tridiagonal form* and it is usually the most expensive among all stages.

(iv) This stage is the core of the chain of transformations. While all the stages above are not data dependent, the solution of the STEVP substantially depends on the distribution of the eigenvalues and it is particularly sensitive to their clustering[6]. Several are the algorithmic choices for this stage. In this chapter we will consider the four most well-known: QR, Bisection & Inverse Iteration (BXINV), Divide & Conquer (D&C), and Multiple Relatively Robust Representations (MRRR). Each one relies on a different strategy so much so that the entire six stages eigensolver inherits its name by the tridiagonal solver used in stage (iv). While this could be common practice it is important to understand that there are also algorithmic variants for the other stages. Consequently several combinations of them are not only possible but quite different. Whatever is the tridiagonal solver of choice the output of this stage are the pairs (λ, z). For this reason stage (iv) is know as *solution of the tridiagonal eigenproblem*. It should be noted in passing that the STEVP also appear as a byproduct of some iterative eigensolvers, most notably the Lanczos method (see Sec. 4.1).

(v) Once we have the pairs (λ, z), it is just a question of tracing back the eigenvectors of STEVP to the eigenvectors of HEVP $y = Qz$ with the first of the so called *back-transformation*. If one was bound to solve just a standard eigenvalue problem, this would be the last stage of the chain.

(vi) Similarly to stage (v), this last stage is meant to compute the eigenvectors of the GHEVP by the *second back-transformation* $x = (L^{\dagger})^{-1}y$ leading to the desired output (λ, x).

Despite the level of complexity of the operations in each stage appears rather low, there are several algorithmic variants for each stage. For example the Cholesky decomposition could be realized in three main variants called respectively right-looking, left-looking and bordered algorithms. Each of these algorithms have distinct performance and memory signatures. Since we cannot cover the fine algorithmic details of all stages we will list, in the following, the most important algorithmic variants of only stages (iii) and (iv). For more details on the algorithmic choices for the rest of the stages we refer to the standard book by Golub and Van Loan [16].

Reduction to tridiagonal form — In general the unitary matrix Q computed at this stage is the composition of a series of projection operators, each one a unitary matrix in itself. The scope of the projection is to eliminate all the entries of C below and above the first lower and upper

[6] A cluster is loosely defined as a set of adjacent eigenvalues densely concentrated around one value with a relative distance substantially smaller than the other neighboring eigenvalues.

sub-diagonals respectively. The two most well-known methods are Givens and Householder transformations.
In the Givens method the matrix Q is the composition of a series of elementary rotation matrices G_{pq} having the only non-zero entries $g_{qq} = g_{pp} = \cos(\phi)$; $g_{ii} = 1$ for $i \neq p, q$, and $g_{pq} = -\bar{g}_{qp} = \rho \sin(\phi)$ with $\rho \in \mathbb{C}$, $\phi \in \mathbb{R}$. By construction G_{pq} is unitary and its action $C^{(1)} = G_{pq}\, C\, G_{pq}^{\dagger}$ can be chosen so as to zero out elements of C with a specific index r, $c_{qr}^{(1)} = c_{rq}^{(1)} = 0$ (the subscript $^{(k)}$ indicates how many elementary transformations were operated on C). For example one can zero out all elements $c_{1,q}\ \forall\ q$ with a sequence of transformations $(p,q) = (2,3), (2,4), \ldots (2,n)$. The resulting matrix $C^{(n-2)} = G_{2n} \ldots G_{2,3}\, C\, G_{23}^{\dagger} \ldots G_{2n}^{\dagger}$ has all elements of the first row and column equal to zero apart from $c_{1,1}^{(n-2)}, c_{1,2}^{(n-2)}, c_{2,1}^{(n-2)}$. Proceeding in the same way one can eliminate all other entries until all is left is just a tridiagonal matrix. Each Givens transformation $C^{(k)} \to C^{(k+1)}$ requires $4(n-r)$ multiplications and for each index r there are $(n-r)$ of them. Summing over all values of $r = 1, \ldots, n-2$ makes the complexity of the entire reduction to tridiagonal form $\sum_r 4(n-r)^2 \approx \frac{4}{3}n^3$. This method is particularly suitable for dense matrices which have some definite structure for the non-zero entries: by avoiding to act on the null elements the Givens method can avoid redundant computations.
The more commonly used Householder method builds Q out of elementary matrices

$$G_k = I - \beta u_k u_k^{\dagger}$$

which are both hermitian and unitary. The vector u_k and constant β are chosen in such way to zero out all the entries of the k column of the matrix $G_k C^{(k)}$ with row indices bigger than $k+1$. By acting on the left with the inverse of G_k (which is the same as the G_k) one eliminates also all entries of the k row with column index bigger than $k+1$. The result is similar to the chain of Givens rotations described above apart from the fact that an Householder transformation achieves this result in a more economical way. In fact the matrix G_k is never used explicitly and only rank two updates are used instead (see Golub Van Loan for details). Because of this property each elementary transformation $C^{(k)} \to C^{(k+1)} = G_k C^{(k)} G_k$ requires only $2(n-k)^2$ operations for a total of $\sum_k 2(n-r)^2 \approx \frac{2}{3}n^3$
A third alternative for the reduction to tridiagonal form was developed relatively recently by Bischof et Al. [17] and goes by the name of two-step successive band reduction (SBR). The basic strategy is to first reduce the dense matrix of the HEVP to a banded form leveraging level 3 BLAS operations (see Sec. 3.2 for a definition of BLAS) and only subsequently reduce the banded matrix to tridiagonal form. Since only the second step uses less performant level 2 and 1 BLAS routines, SBR shows better performance with respect to the classic Householder method. On the downside SBR needs a total of $\frac{4}{3}n^3$ operations for the reduction and $2n^3$ for the accumulation of the matrix Q on the fly. This implies that SBR is a very convenient method when one is interested in only the eigenvalues, since they don't need the accumulation of the similarity matrix Q, while it may be penalizing if also the eigenvectors are required.

Tridiagonal eigensolver — Four are the main algorithms that are used to solve for the tridiagonal problem. They are distinguished by the solving strategy, the memory requirements, complexity and the ability to solve for just a subset of the spectrum. In the following we provide a short description and indicate which are the most suitable for DFT computations.
QR uses a series of similarity transformations, preserving the tridiagonal structure of T, which turn progressively off the sub-diagonal elements. The algorithm achieves this target by a so

called bulge-chasing procedure which includes implicit shifts and deflation techniques. The total complexity of the algorithm is $\mathcal{O}(sn^3)$, where s indicates the median number of bulge chases per eigenvalue. The QR algorithm cannot be used to solve for subset of the eigenspectrum and so it is not particularly suitable for DFT computations.
BXINV is the first of the two algorithms which is capable of solving for a subset of the eigenpairs. The algorithm uses Sturm sequences (bisection algorithm) with a total of $\mathcal{O}(kn)$ operations for k eigenvalues. If such eigenvalues are well separated then BXINV requires another $\mathcal{O}(kn)$ to compute the relative eigenvectors (inverse iteration method). In the case the eigenvalues are grouped in tight clusters the complexity can grow up to $\mathcal{O}(n^3)$. The grow in complexity is motivated by the need of re-orthogonalize eigenvectors since numerical orthogonality is not automatically guaranteed for clustered eigenvalues. This algorithm is probably the one currently most used for DFT computations involving dense eigenproblems. As we will see later BXINV is by far outperformed by its close cousin MRRR which should be preferred to it.
D&C strategy is rather well described by its name. This algorithm decomposes T in a hierarchical tree of smaller submatrices and rank-one updates. At the bottom of the tree each child submatrix is solved using a secular equation and the process is repeated going from child to father until the tree is complete. D&C cannot compute subset of the eigenspectrum and its complexity is $\mathcal{O}(n^3)$. The complexity can often be reduced substantially by a deflation process when certain entries in the eigenvectors of the submatrices are small enough. Despite being not optimal for DFT computations, D&C can be quite performant so as to be used in DFT computations provided one discards, in the end, the part of the spectrum not required.
MRRR is the second of the algorithms capable of computing for a subset of the spectrum. This algorithm is a sophisticated variation of BXINV which avoids altogether the re-orthogonalization of the eigenvectors. Consequently MRRR complexity is approximately reduced to $\mathcal{O}(n^2)$. In practice the overall complexity depends on the clustering of the eigenvalues. Due to its low complexity and the ability to solve for a portion of the eigenspectrum without the need for costly orthogonalizations, this is the most indicated algorithm for DFT computations and should almost always be preferred to BXINV.

3.2 Libraries, performance and parallelism

Since numerical linear algebra deals with vectors and matrices the most common and important operations are included in specialized and optimized libraries. Among the most well-known are the Basic Linear Algebra Sub-routines (BLAS) and the Linear Algebra PACKage (LAPACK). One of the most important practice that can improve the performance of an algorithm-derived routine is the correct use of the kernels already present in these standard libraries.

BLAS and LAPACK — The main motivations behind the Basic Linear Algebra Subroutines were modularity, efficiency, and portability through standardization [18]. The BLAS library consists of three levels, corresponding to routines for vector-vector, matrix-vector and matrix-matrix operations [19, 20]. From a mathematical perspective, it might appear that this structure introduces unnecessary duplication: For instance, a matrix-matrix multiplication (a level 3 routine) can be expressed in terms of matrix-vector products (level 2), which in turn can be expressed in terms of inner products (level 1). The layered structure is motivated by the increased efficiency of level 2 and 3 routines on architectures with a hierarchical memory. In fact BLAS 1, 2, and 3 are capable respectively of 1/2, 2 and n/2 operation counts per number of memory accesses, giving the higher-level routines a better opportunity to amortize the costly memory

accesses with calculations. With respect to the full potential of a processor, the efficiency of BLAS 1, 2, and 3 routines is roughly 5%, 20% and 90+%, respectively. Moreover, the scalability of BLAS 1 and 2 kernels is rather limited, while that of BLAS 3 is typically close to perfect. In practice, this means that level 3 routines attain the best performance and should be preferred whenever possible. In other words BLAS succeeded in providing a portability layer between the computing architecture and both numerical libraries and simulation codes.
In addition to the reference library,[7] nowadays many implementations of BLAS exist, including hand-tuned [21, 22], automatically-tuned [23], and versions developed by processors manufacturers (Intel-MKL, IBM-ESSL, AMD-CML). More importantly BLAS kernels are heavily used in most of standard libraries (e.g. LAPACK, ScaLAPACK, etc.) implementing the linear algebra operations we have described in the six stages of the GHEVP.
The LAPACK library in its modern form already includes BLAS kernels and add, on top of that, many other routines covering almost the entire spectrum of standard linear algebra operations. For example we can find in it routines to solve triangular linear systems as well as Cholesky decomposition used respectively in stage (vi) and stage (i) of the GHEVP. This library was built in the '90s on top of other library packages developed 20 years earlier like EISPACK. In LAPACK one important section is devoted to the solution of the Symmetric Eigenvalue Problem (which includes also the complex case). Routines for the solution of all three type of eigenproblems are included allowing for several choices depending, for example, on the number of eigenpairs required as well as the storing of the matrices in memory.

Tridiagonal eigensolvers on a single core — We now illustrate how some of the tools of Section 2 can be used to analyze the performance of the tridiagonal eigensolvers described in the previous subsection. Results illustrated below were conducted on a single processor using the LAPACK eigensolver implementations. Since we deal with a specific library, we refer to each eigensolver interchangeably by its acronym or the name of its relative routine as outlined in Table 1.

Table 1: *LAPACK tridiagonal eigensolvers*

Algorithm	Routine	Workspace	Eigenpairs subsets
QR	`STEV`	Real: $2n-2$	No
BXINV	`STEVX`	Real: $8n$	Yes
D&C	`STEVD`	Real: $1+4n+n^2$	No
MRRR	`STEVR`	Real: $18n$	Yes

Experimental tests using both artificial and practical matrices show without any doubt that `STEVR` and `STEVD` are typically much faster than `STEV` and `STEVX` [24]. On average MRRR is the fastest among the four algorithms with some exceptions in special cases. These conclusions are supported by an experimental measure of the "effective" complexity of each algorithm. By comparing the CPU time to completion and counting the total number of operations performed, it is possible to infer an experimental complexity value based on the assumption that the performance of each solver does not change with the size of the problem examined.
By testing on a large set of matrices it results that the effective complexity of `STEV` is $\mathcal{O}(n^{2.9})$, very similar to the one for `STEVX`. For `STEVR` the effective complexity is $\mathcal{O}(n^{2.2})$, a sligtly

[7] Available at `http://www.netlib.org/blas/`.

higher value than the theoretical one. STEVD instead registers a lower value, $\mathcal{O}(n^{2.8})$, respect to the theoretical one. So from the operation count point of view the higher performance of MRRR is quite justified.
The results for QR and BXINV are in line with the fact that neither the bulge chasing part of QR nor the bisection and inverse iteration of BXINV use any routine from the levels 2 or 3 of BLAS. The interesting piece of data is that, while MRRR execute many less operations than the other algorithms, it does so at a higher cost. In other words the number of divisions STEVR performs is always a significant fraction of the total number of operations. Since divisions cost up to 4 times more than multiplications, the measured performance is lower than expected resulting in a higher effective complexity. STEVD experiences the opposite outcome due to the combined effect of using BLAS 3 to update the eigenvector matrix, together with the mechanism of deflation. If there is a lot of deflation STEVD performs many fewer scalar operations (slow), while if there is little deflation most of the flops are performed by calls to level 3 BLAS (fast). Overall STEVD experience a speed-up respect to its theoretical complexity.
The above example illustrate how performance for an algorithm depends on several factors including complexity, block operations[8], efficiency of specialized kernels, etc. To conclude let me also remind that the performance depends also on the hierarchy of the cache memories of the computing architecture. For example STEVD may well be the fastest routine for a specific class of eigenproblems (e.g Wilkinson matrices) but the same routine uses up to $\mathcal{O}(n^2)$ workspace so that on some specialized machines such routine may not be used for large problems anymore. In the latter case the lower workspace ($\mathcal{O}(n)$) favor STEVR the optimal algorithm of choice.

ScaLAPACK, ELPA and Elemental — Where LAPACK is the most well-known library for dense numerical linear algebra single processor routines[9], ScaLAPACK is probably the most renown library for heterogeneous computing platforms. It is a adaptation of most LAPACK routines to distributed memory architectures using the Message Passing Interface (MPI) protocol. It requires PBLAS, the parallel version of BLAS, which is currently included in the library. A specific section of this library is devoted to the solution of dense eigenproblems. ScaLAPACK was also the first library using block cyclic data distribution for dense matrices together with block-partitioned algorithms. This was an important design feature introducing a tension between having block sizes that are large enough for local BLAS efficiency yet small enough to avoid inefficiency due to load-imbalance. In ScaLAPACK this tension is intrinsically included in the design by linking the distribution block sizes to the algorithmic block sizes. The library is written in Fortran and it is structured in low-level modular routines which follow the same pattern of the LAPACK ones.
As can be seen from table 2 ScaLAPACK includes only one routine for GHEVP which uses BXINV for the tridiagonal solver while all four algorithms are represented for the HEVP. This is actually not a limitation of ScaLAPACK since one can combine routines for the other stages other than (iv) and build one's own GHEVP solver. For example it was noted in [25] that one should always use the PZHENTRD for the reduction to tridiagonal form (for a square grid of processes) instead of PZHETRD: the former, which is designed for a rectangular grid of processes, is so much less performant than it is always convenient to use less cores arranged on a square grid.
The are modern alternatives to ScaLAPACK routines for dense eigensolvers. In particular in re-

[8] conceptually similar to algorithmic block size used in parallel libraries
[9] Recently some LAPACK libraries like MKL provide also multi-threaded functionalities

Table 2: *ScaLAPACK eigensolvers routines*

Algorithm	Routine (Complex)	Routine (Real)	Eigenproblem type
BXINV	PZHEGVX	PDSYGVX	GHEVP
QR	PZHEEV	PDSYEVX	HEVP
BXINV	PZHEEV	PDSYEV	HEVP
D&C	PZHEEVD	PDSYEVD	HEVP
MRRR	PZHEEVR	PDSYEVR	HEVP
D&C	–	PDSTEVD	STEVP
MRRR	–	PDSTEVR	STEVP
BXINV	–	PDSTEBZ	STEVP eigenvalues only
BXINV	PZSTEIN	PDSTEIN	STEVP eigenvectors only

cent years there has been quite a large interest in exploring new framework for parallel libraries for large dense eigenproblems. Among the several attempts in this direction the two most recent and performant are ELPA and Elemental. While ELPA is a set of routines specifically designed for eigenvalue problems, Elemental is a new complete framework for dense linear algebra operations which also include a driver for dense eigenproblems.

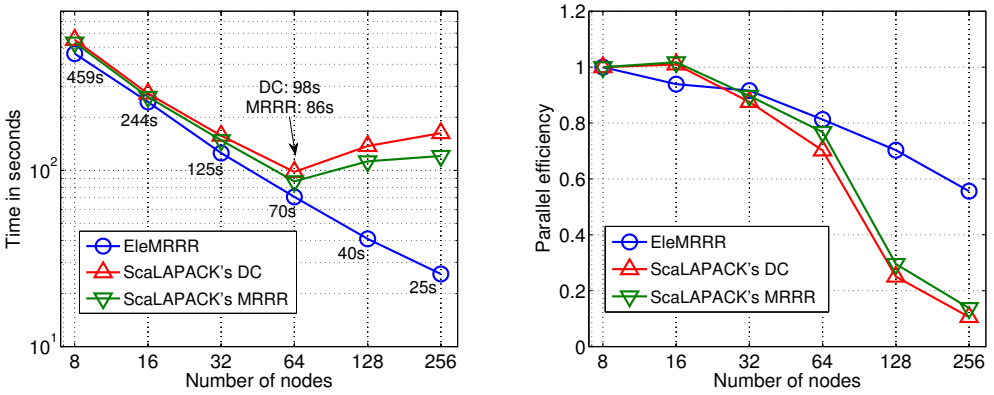
ELPA is a set of Fortran subroutines (modules) which can be compiled with an application of choice or as separate library which can also be linked to from C, or C++ code [26]. ELPA is based on the ScaLAPACK framework and cannot be used independently from it. From this point of view the ELPA library builds on top of ScaLAPACK and it is not, strictly speaking, an alternative to it. As such, necessary prerequisite libraries for ELPA include BLAS, LAPACK, Basic linear algebra communication subroutines (BLACS), and obviously ScaLAPACK. ELPA is an MPI only implementation; there are no hybrid parallelization (MPI/OpenMP) available. Consequently the ELPA library works both on a single-node, shared memory environment, as well as on large clusters of distributed memory nodes.

The library provides modular routines for two different approaches to solve large eigenvalue problems: the ELPA 1 STAGE modules, including the routine solve_evp_real , focus on reducing communication overheads and maximizing cache performance for existing standard ScaLAPACK routines (PDSYEVR). ELPA 2 STAGE routines such as solve_evp_real_2stage add extra steps into the traditional three-step HEVP approach outlined in the previous section. An intermediate banded representation of the matrix is formed during the reduction to tridiagonal form using a variant of the SBR algorithm. Likewise an intermediate banded form is generated during the back transformation stage of the calculation. The 2 Stage routines provide the real novelty of the ELPA library. Both ELPA solvers can be directed to calculate a subset of eigenpairs if preferred by the user.

Elemental is a modern framework for distributed memory dense linear algebra [27] which is designed to overcome the possible problem deriving from linking the algorithmic block size to the distribution block sizes. Elemental restricts the algorithmic block sizes to $m_b = n_b = 1$. The core of the library is the two-dimensional cyclic element-wise ("elemental" or "torus-wrap") matrix distribution. This approach is based on the observation [21] that the optimal algorithmic block size should be related to the square root of the size of the L2 cache memory. The Elemental approach for the distribution grid avoid linking the distribution with the algorithmic block sizes and so eliminates the issue between load balancing and filling the L2 cache. For a given

number $p > 1$ of processors there are several possible choices for r and c forming different grid shapes $(r, c) \doteq r \times c$. Since the grid shape can have a significant impact on the overall performance, careful experiments should be undertaken in order to determine the best choice of (r, c). Elemental parallel eigensolver is based on a parallelized version of MRRR [28, 25], namely EleMRRR.

Scalability and efficiency of parallel libraries — Implementing a scalable eigensolver implies that all the stages, even the less expensive ones, should have the same degree of scalability. We will briefly illustrate how scalability depends on many factors and may vary as a function of the size of the problem as well as the number of cores utilized. In the following we will show some example of scalability for ScaLAPACK and Elemental. The numerical results and the plots are extracted from the paper [25] by generous concession of the authors.



(a) Execution time (b) Parallel efficiency

Fig. 1: *Strong scalability for the computation of all eigenpairs with matrices of size 20,000. Red and green lines corresponds to ScaLAPACK eigensolvers based on the tridiagonal D&C and MRRR. The blue line corresponds to the parallel MRRR-based eigensolver implemented in Elemental. (By courtesy of Petschow and Bientinesi)*

The strong scaling plots shows that there is not much difference between the three parallel eigensolvers up to 512 cores. At that point ScaLAPACK solvers performance degrades dramatically due to the MPI communication design decision: while the elemental distribution of Elemental does not force an algorithmic block, ScaLAPACK blocked communication hamper the scalability of the algorithm implementation over a certain number of cores. This is more evident by looking at the parallel efficiency which degrades quite dramatically for ScaLAPACK solvers above 512 cores, where the reference point is chosen to be 64 cores. Observe the perfect efficiency of `PZHEEVD` and `PZHEEVR` for 16 nodes (equivalent to 128 cores) due to the use of the very efficient ScaLAPACK `PZHENTRD` routine. The plots for weak scalability shows how Elemental scales much better for larger size eigenproblems than either ScaLAPACK D&C and MRRR. This result implies an almost perfect parallel efficiency and, consequently the possibility to solve increasingly bigger eigenproblems in a still reasonable amount of time.
The superior performance of the Elemental eigensolver (EleMRRR) is due to two main factors: 1) the elemental distribution strategy among processes favor a more fine coarse communication pattern which is close to optimal and 2) the tridiagonal eigensolver has been parallelized using a dynamic allocation of the load among all processes. The latter characteristic allows EleMRRR

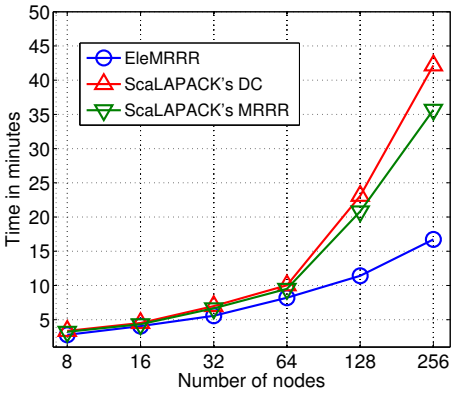


Fig. 2: *Weak scalability for the computation of all eigenpairs with matrices of increasing size. Red and green lines corresponds to ScaLAPACK eigensolvers based on the tridiagonal D&C and MRRR. The blue line corresponds to the parallel MRRR-based eigensolver implemented in Elemental. (By courtesy of Petschow and Bientinesi)*

to scale also when using thousands of cores. The same cannot be said for ScaLAPACK D&C tridiagonal stage: the fraction of time spent on this stage by the eigensolver goes from 4.5% on 64 cores to 41% on 2048 cores determining its drop in performance.

It is important to point out that both ELPA and EleMRRR are quite recent additions and are still not used by the majority of DFT codes dealing with dense eigenvalue problems. On the opposite the use of the more established ScaLAPCK is almost ubiquitous. This seem quite in contradiction with the desire to simulate physical systems with a higher number of atoms, which in turn needs a more efficient and scalable eigensolver. The main reason behind the lack of change is to be found in the need of a profound change for the data distribution pattern. From this point of view ELPA requires much less man-programming hours than Elemental. Despite this increased handicap, future DFT codes, which want to efficiently run on massive parallel architectures, will inevitably have to step up and face the initial investment.

4 Iterative eigensolvers

As mentioned at the beginning of section 3, for a sparse eigenproblem it is advisable to use an eigensolver which preserves the structure of the sparsely populated matrices which define it. It is important to notice that a matrix is defined as sparse when the number of non-zero entries (nnz) is less than a few percentage points of n^2. This is particularly important because these matrices are usually stored in memory in sparse format: only non-zero entries and their indices are specified. Consequently saving in memory only the nnz data frees up quite a bit of memory. Iterative methods do not attempt to directly diagonalize the matrix but instead strive at determining the eigenspace (or a subspace of the eigenspace) of the problem. This result is attained by repeadetly multiplying one (many) trial vector(s) with the matrix defining the HEVP[10] and in some way aligning it to the dominant eigenvector (eigenspace). This idea is at the base of the so-called power method and provides the basic principle on top of which all other iterative methods are built.

[10]The GHEVP is usually treated by reducing it to standard form even if there are several possible alternatives.

The power method relies on the basic concept that any random vector v can be decomposed as a linear combination $v = \sum_j \gamma_j x_j$ of the eigenvectors basis set $\{x_1, \ldots, x_n\}$ with $|\lambda_1| > |\lambda_2| > |\lambda_3| > \ldots$. The repeated multiplication of A on v from the left results in

$$Av = \sum_j \lambda_j \gamma_j x_j \quad \Longrightarrow \quad v^{(k)} = A^k v = \sum_j \lambda_j^k \gamma_j x_j = \lambda_1^k \left[x_1 + \sum_{j \geq 2} \left(\frac{\lambda_j}{\lambda_1} \right)^k \gamma_j x_j \right].$$

For large enough k the eigenvector with largest eigenvalue dominates over the others and $v^{(k)}$ converge to x_1 at the rate with which all the coefficients $\left| \frac{\lambda_j}{\lambda_1} \right|^k$ become negligible.
Since iterative methods are all roughly based on the power method, their effectiveness is grounded in multiple repetitions of matrix-vector multiplications. As such, iterative methods work on the vectors and not on the matrix, thus maintaining intact its sparse structure. While preserving the matrix sparse structure, efficient matrix-vector multiplications cannot rely on BLAS and need to be optimized for the specific problem or class of problems. This characteristic makes very difficult to know a priori the complexity of an iterative eigensolver. This is the more so since one does not know in advance when a trial vector would converge to an eigenvector.
In theory, there is nothing that prevents iterative methods to be used on dense matrices. In fact if the dense matrix is not too large and the desired fraction of the spectrum is very small ($< 1\%$), it is common belief that iterative methods could still be competitive, performance-wise, with direct ones. As long as the number s of matrix-vector multiplications required for convergence of the residuals is less than the inverse of the fraction of spectrum desired f, the iterative solver complexity $\mathcal{O}(s * f * n^2) \lesssim \mathcal{O}(n^3)$. In other words the iterative solver is no more complex than a direct one. We will see that this fact can be exploited in the case of some tailored eigensolvers, an example of which is described in Sec. 5.
Another issue to address with iterative solvers is the marked distinction between solving for the GHEVP and the HEVP. This difference has its source in the lack of a clear path which takes the GHEVP and transforms it to an HEVP. In general one would like to preserve the structure of the matrix B and so avoid expensive factorizations which would inevitably lead to a dense factor L and spoil the sparsity of the problem. Avoiding factorizations depends very much on how well-conditioned is the B matrix. If the condition number is low one could "invert" B and solve for $B^{-1}Ax = \lambda x$. On the contrary if $\kappa(B)$ is quite large, factorizations are unavoidable and the first two stages of the direct method are used to reduce the problem to standard form. In these cases it is customary to rely on incomplete factorization which limit the density of L and may still produce good enough solutions. The rest of this section will not enter in the merit of describing how to solve GHEVP and exclusively deal with solving HEVP with iterative eigensolvers.
The large class of iterative eigensolvers can be divided in 3 major groups. Subspace iteration based methods, Krylov-based methods, Davidson-based methods.

4.1 Three classes of iterative methods

Subspace iteration methods — Subspace iteration is just a generalization of the power method. Instead of iterating the multiplication of A with v ($v^{(i)} = Av^{(i-1)}$) one can iterate on a space of dimension larger than one $V \in \mathbb{C}^{n \times k}$. Iterating on such initial space would, in general, lead to a *dominant* k-dimensional subspace $\mathcal{U}$ associated with the first k eigenvalues $|\lambda_1| \geq \cdots \geq |\lambda_k|$. The global convergence ratio now clearly depends on the magnitude of the largest among the coefficients required to be negligible $\left| \frac{\lambda_{k+1}}{\lambda_k} \right|$.

In practice, if one use subspace iteration in its purest form $V^{(i)} = AV^{(i-1)}$, each single $v_j^{(i)} \in V^{(i)}$ becomes mostly aligned with the eigenvector corresponding to $\lambda_{\max}$. After a sufficient number of iterations the subspace could become rank-deficient leading to multiple copies of the same eigenvectors. In order to avoid this undesirable effect, subspace iteration is usually used in combination with a shift $A - \sigma I$. The dominant eigenvector of the shifted matrix would be the one corresponding to the eigenvalue closer to σ. By using a different shift at each iteration distinct eigenspaces can be enhanced at each iteration

$$V^{(i)} = (A - \sigma_i I) \cdots (A - \sigma_1 I)V$$

To work properly this method assumes the a priori knowledge of shifts $\{\sigma_1, \ldots, \sigma_k\}$ close enough to the true eigenvalues. Such knowledge is not often available but can be inferred by evaluating the Ritz values after a certain number of non-shifted iterations.
The procedure described above is equivalent to using for a single iteration a polynomial $p(t) = (t - \sigma_i) \cdots (t - \sigma_1)$ having zeros close to the k dominant eigenvalues. In addition to generalize the subspace iteration to polynomials one can orthonormalize the resulting vectors using the Gram-Schmidt procedure

$$p(A)V^{(i-1)} = [p(A)v_1^{(i-1)}, \ldots, p(A)v_k^{(i-1)}] \longrightarrow [q_1^{(i)}, \ldots, q_k^{(i)}] = Q^{(i)} \text{ with } (Q^{(i)})^\dagger Q^{(i)} = I.$$

The convergence to the solutions of the iterates can be tricky. In general computing the error (as the distance between successive vectors) may not guarantee that when the error is small we reached convergence. A better criterion is to check the residuals for the approximated eigenvectors and declare the eigenpairs converged when residuals are below the required tolerance. In general one measure the rate of convergence of the process by looking at the difference between the residuals across two iterations. If this difference decrease linearly one refers to it as linear convergence. It is a known results of numerical linear algebra that subspace iteration can at best converge linearly.

Krylov subspace methods — Krylov-based methods refer to a specifically constructed sequence of subspaces approximating an invariant subspace of the entire spectrum. The relative space is called Krylov and is indicated by $\mathcal{K}_k(A, v)$. The ingenuity of the Krylov method is in building a sequence of subspaces of increasing dimensionality (as opposed to the fixed dimensionality of subspace iteration). In practice the subspace is characterized directly by an orthonormal space which is the by-product of the construction.
One start with a normalized vector $u = u_1 \equiv \mathcal{K}_1(A, u)$ and compute u_2 as

$$u_2 h_{2,1} = Au_1 - h_{1,1}u_1 \quad ; \quad \mathcal{K}_2(A, u) = \text{span}\{u_1, u_2\}$$

where $h_{1,1}$ is chosen so as to make u_1 orthogonal to u_2, and $h_{2,1}$ is just a scale factor. The whole process is then repeated j times to find $\mathcal{K}_j(A, u) = \text{span}\{u_1, \ldots, u_j\}$. At the $j + 1$ step, one first checks if $Au_j \in \text{span}\{u_1, \ldots, u_j\}$. If this is the case, the coefficent $h_{i,j}$ should be chosen such that $Au_j = \sum_{i=1}^{j} u_i h_{i,j}$. In this way $h_{j+1,j} = 0$ and u_{j+1} is indeterminate. At this point the Krylov process terminates and the resulting subspace is invariant.
In general the k-step of the Krylov process can be represented in matricial form as

$$AU_k = U_k H_{k+1,k}$$

where the matrix U_k collects all the u vectors, and $H_{k+1,k}$ is a matrix accumulating all $h_{i,j}$. If one defines the upper Hessenberg matrix deleting the bottom row of $H_{k+1,k}$ then the process is better described by

$$AU_k = U_k H_k + u_{k+1} h_{k+1,k} e_k^T.$$

Whenever $h_{k+1,k} = 0$ then $A = UHU^{-1}$ which clearly shows that the matrix H corresponds to the eigenproblem A reduced to the invariant subspace $\mathcal{K}_k(A, u)$. Even when the Krylov process is only partially completed it produces a quasi-similarity transformation.
Thus even if $h_{k+1,k}$ is not equal to zero, the Krylov process leads to a subspace which contains good approximants to the eigenvectors of some of the peripheral eigenvalues of A. This property constitutes the basis of the the so-called Arnoldi method of *implicit restarts*. It has been successfully implemented by Sorensen et Al. in the package ARPACK [11]. The aim of the Arnoldi procedure is to keep $m \ll k$ vectors of the Krylov subspace which are rich in the components of the good approximants to the eigenvectors, discard the rest and start the Krylov process again from these m vectors. This is achieved through a GR process whose details are not important for the purpose of this chapter. We remit the reader to the vast literature on the subject [29].

Algorithm 1 Symmetric Lanczos algorithm with re-orthogonalization

Require: Matrix A of the HEVP and initial vector u
Ensure: NEV wanted eigenpairs (λ, x).

$u_1 \leftarrow u/\|u\|$
$U_1 \leftarrow [u_1]$
for $j = 1, 2, 3, \ldots \rightarrow$ NEV (λ, x) **do**
  $u_{j+1} \leftarrow A u_j$
  $\alpha_j \leftarrow u_j^\dagger u_{j+1}$
  $u_{j+1} \leftarrow u_{j+1} - \alpha_j u_j$
  if $j > 1$ **then**
    $u_{j+1} \leftarrow u_{j+1} - \beta_{j-1} u_{j-1}$
  end if
  $\delta_{1:j} \leftarrow U_j^\dagger u_{j+1}$
  $u_{j+1} \leftarrow u_{j+1} - U_j \delta_{1:j}$ ▷ RE-ORTHOGONALIZATION
  $\alpha_j \leftarrow \alpha_j + \delta_j$
  $\beta_j \leftarrow \|u_{j+1}\|$
  if $\beta_j = 0$ **then**
    Flag span$\{u_1, \ldots, u_j\}$ is invariant
    Exit
  end if
  $u_{j+1} \leftarrow u_{j+1}/\beta_j$
  $U_{j+1} \leftarrow [U_j \; u_{j+1}]$
end for

Since the Krylov process preserves the symmetry of the matrix A, in the case of the Hermitian (or symmetric) eigenvalue problem the matrix H_k is actually tridiagonal. The net result is that the whole process of building H is actually reduced to a 3-term recurrence relation

$$u_{j+1}\beta_{j+1} = Au_j - \alpha u_j - \beta_{j-1} u_{j-1}$$

In this case the Krylov-Arnoldi process is called Lanczos process (see Alg. 1)

The Krylov process is an instance of the the Gram-Schmidt method and as such is vulnerable to roundoff errors. Over the course of several steps executed in floating point arithmetic, the orthogonality of the vectors steadily deteriorate. The standard remedy, as suggested by Kahan in an unpublished work, is to orthogonalize twice. This practice has also been confirmed by numerical analysis [30]. Thus the safest way to preserve orthogonality is to save all the vectors and orthogonalize against them.
By re-orthogonalizing, the Lanczos process looks very similar to the Arnoldi, so one may wonder why bothering with the 3-term recurrence relation. The simple answer is that the Lanczos process has the advantage of preserving the structure of the matrix H leaving it symmetric like the original A. Then the restart can be performed through a symmetric QR algorithm which also keep the structure intact and has a lower complexity than the full GR.

Davidson methods — The original Davidson method was devised by the homonymous author to compute the lowest energy levels and corresponding wave functions for eigenvalues problems arising in quantum chemistry. Davidson algorithm builds a subspace of increasing dimensionality by adding a new vector at each step in a way similar to the Lanczos algorithm. The main difference lies in the choice of vector: instead of being extracted by Au_{k-1} the additional vector is obtained with a correction equation for the residual $r_{k-1} = (Au_{k-1} - \hat{\lambda}_{k-1}u_{k-1})$ sometimes also referred as diagonal *preconditioning step*.
Several are the variants of the Davidson method, some of which are also implemented in a block version. All these can be classified in two major groups: Generalized Davidson (GD) and Jacobi-Davidson (JD) methods. In its original and simplest form the algorithm would look only for the largest or smallest eigenpair. In Alg. 2 we present a rather simple formulation which shows the major differences with Lanczos are in lines 5 and 12. The simplest Davidson method

Algorithm 2 Davidson algorithm

Require: Matrix A of the HEVP and initial vector u
Ensure: largest eigenpair $(\lambda_{\max}, x)$.

1: $u_1 \leftarrow u/\|u\|$
2: $U_1 \leftarrow [u_1]$
3: **for** $j = 1, 2, 3, \ldots \rightarrow$ NEV (λ, x) **do**
4: $H_j \leftarrow U_j^\dagger A U_j$ ▷ RAYLEIGH-RITZ QUOTIENT
5: Compute largest eigenpair $(\hat{\lambda}_j, \hat{y}_j)$ of H_j
6: Compute Ritz vector $x_j \leftarrow U_j y_j$
7: Compute the residual $r_j \leftarrow (\hat{\lambda}_j I - A)x_j$
8: **if** $r_j <$ TOL **then**
9: $(\lambda_{\max}, x) \leftarrow (\hat{\lambda}_j, \hat{y}_j)$
10: Exit
11: **end if**
12: Correction equation $t_{j+1} \leftarrow (\hat{\lambda}_j I - D)^{-1} r_j$
13: $\delta_{1:j} \leftarrow U_j^\dagger t_{j+1}$
14: $t_{j+1} \leftarrow t_{j+1} - U_j \delta_{1:j}$ ▷ ORTHOGONALIZATION
15: $t_{j+1} \leftarrow t_{j+1}/\|t_{j+1}\|$
16: $U_{j+1} \leftarrow [U_j\ t_{j+1}]$
17: **end for**

does not rely on any kind of restart but keeps building a subspace and then constructing the

Rayleigh-Ritz quotient out of which it extracts an approximant to the largest eigenpair. The subspace is incremented by solving a "preconditioned" linear system associated with the HEVP where the known term is the residual of the approximant and D is the main diagonal of A.
Davidson method can be straightforwardly generalized to a block implementation where the U_j is incremented with a block of vectors and a block of eigenpairs is sought after. The block adaptation is particularly relevant when the resulting code are ported over parallel computing architectures. What is not obvious is how to generalize the correction equation on line 12. In general this equation can be written as $C_{j+1}t_{j+1} = -r_j$ and solved with a low level of accuracy. The operator C_{j+1} can be as simple as $(A-\hat{\lambda}_j I)$ or as complicated as $(I-x_j x_j^\dagger)(A-\hat{\lambda}_j I)(I-x_j x_j^\dagger)$. The latter choice is at the base of the JD method: by solving the correction equation orthogonally to the Ritz vector x_j, JD avoids the well-known effect of *stagnation*. In addition to an improved correction equation a preconditioner can be used. In the case of JD the inverse of such preconditioner needs to be inverted orthogonally to x_j. Additionally the Davidson methods could be restarted for efficient use of the memory and better convergence of the subspace approximants.
It is worth to mention an algorithm which does not quite fit the above grouping: the Locally Optimal Block Preconditioned Conjugate Gradient method (LOBPCG). Similar to a block version of GD the LOBPCG can deal directly with GHEVP. Developed in 2001 by Knyazev [31, 32], this algorithm uses a locally optimized version of a three-term recurrence relation for the preconditioned conjugate gradient method. In practice the Rayleigh-Ritz method is used for the eigenpencil on a trial subspace generated by the current guess for the Ritz vector, the preconditioned residual, and a third Ritz vector built by maximizing the Rayleigh quotient. Knyazev implemented a block version of the algorithm where the three-term relation is generalized for a block of vectors. LOBPCG can deal directly with GHEVP only if B is well-conditioned. When $\kappa(B)$ has a large value the performance of the algorithm deteriorates quite rapidly and reduction to standard form is more stable [33].

4.2 Libraries

ARPACK and PARPACK — The first and foremost best known library of iterative eigensolver is the Arnoldi Restarted Package (ARPACK). This library is based upon an algorithmic variant of the Implicit Restarted Arnoldi Method (IRAM) which reduces to the Implicit Restarted Lanczos Method (IRLM) for Hermitian or symmetric eigenproblems. Both variants are just instances of the general Implicitly Shifted QR technique applied to the the Krylov-Arnoldi process. ARPACK is conceived to solve for large scale symmetric, or non-symmetric eigenvalue problems. The software is designed for sparse and structured matrices such that the matrix-vector product Au would involve only $\mathcal{O}(n)$ operations instead of the standard $\mathcal{O}(n^2)$.
The collection of subroutines making up the library are written in Fortran77. For many standard problems, there is no need for an explicit matrix factorization. Only the action of the matrix A on a vector is required. The software is aimed at computing a small fraction of the total number of eigenvalues with either largest real part or largest magnitude. Storage requirements are on the order of $n \times s$ with s being the number of required eigenpairs. No extra auxiliary storage is required. A set of Schur basis vectors for the desired s-dimensional eigenspace is computed which is numerically orthogonal to working precision. Numerically accurate eigenvectors can be computed on request.
The PARPACK software package realizes the parallelization of the ARPACK library for distributed memory architectures. It has been designed so as to provide the user with a Single Pro-

gram Multiple Data (SPMD) template. The reverse communication interface, which is the most important design feature of ARPACK, has motivated the parallelization strategy. The interface enables PARPACK to be internally parallelized by avoiding to impose predetermined parallel decompositions on A and on the user-provided matrix-vector product. The call to PARPACK preserve the same structure as ARPACK, the only difference consisting in the inclusion of the Basic Linear Algebra Communication Subprograms context. The net result is outlined in 3 main steps: 1) replicating H_j on every processor, 2) distributing (blocked by rows) U_j on a 1-dimensional processor grid, 3) distributing the workspace accordingly. The greater part of the communication takes place during the Gram-Schmidt orthogonalization and possibly in the user supplied matrix-vector multiplication. Clearly the intrinsic serial nature of the algorithm does not lend itself to parallelization on massive number of processors when the parallel efficiency degrades substantially.

SLEPc — The Scalable Library for Eigenvalue Problem Computations (SLEPc) [34], is a software library for the solution of large sparse eigenproblems on parallel computing architectures. The majority of SLEPc routines are intended to be used for the solution of both generalized and standard eigenvalue problems in their linear or non-linear version. SLEPc provides a large set of different methods and focus on the choice of the most appropriate one in relation to the application which generated it. Most of the methods offered by the library are indeed subspace methods. The default eigensolver is a slight modification of the classic Krylov process named Krylov-Schur method [35, 36]. The library also offers routines implementing JD and Conjugate Gradient methods as well as an interface to several external packages like ARPACK, BLZPACK, TRLAN, BLOPEX, and PRIMME. The library does not limit itself to the solution of Hermitian problems but it extends also to non-symmetric and more general complex-valued eigenproblems. Among its other functionalities, it provides routines for Singular Value Decompositions and tools of spectral transformations such as shift and invert, etc..
SLEPc is built on top of the Portable Extensible Toolkit for Scientific Computation (PETSc) [37]. As an extension to PETSc toolkit, SLEPc inherits a variety of tuned data structures, multivector operations, matrix-vector and preconditioning operators, but it cannot run as stand-alone with applications that do not use PETSc. This means that PETSc must be previously installed in order to use SLEPc. This library enforces the same programming paradigm as PETSc making it quite easy to manipulate for users already acquainted with the use of the latter. Users who are not familiar with PETSc are invited to first get accustomed to the basic concepts of such framework before endeavor in the use of SLEPc.

PRIMME and BLOPEX — The PReconditioned Iterative Multi Method Eigensolver (PRIMME) [38] is a software package for the solution of large, sparse Hermitian and real symmetric standard eigenvalue problems. PRIMME constitutes a significant effort towards the realization of a robust and effective code for the solution of large and difficult eigenproblems when matrix factorization is not a viable option and the user has to rely only on matrix-vector operations and, possibly, a preconditioning operator. PRIMME developers stress that the library design strategy is to provide a flexible eigensolver with the following salient characteristics: 1) preconditioned eigen-methods converging near optimally under limited memory, 2) the maximum robustness possible without matrix factorization, 3) flexibility in mixing and matching among most currently known features, 4) efficiency for most architectural layout, and 5) a friendly user interface that requires no parameter setting from end-users but permits full ex-

perimentation by experts. The algorithmic framework of PRIMME is built on top of the two near optimal methods derived from the class of Davidson algorithms: GD+k and JDQMR. It is remarkable that these two algorithms also provide the structure to show how other algorithms can be parameterized within this framework.
The Block Locally Optimal Preconditioned Eigenvalue Xolvers (BLOPEX) [32] is a package, written in C and MATLAB/OCTAVE, that includes an eigensolver implemented with the Locally Optimal Block Preconditioned Conjugate Gradient Method. Its prominent characteristics are: 1) a matrix-free iterative method for computing several extreme eigenpairs of symmetric positive generalized eigenproblems, 2) a user-defined symmetric positive preconditioner, 3) robustness with respect to random initial approximations, variable preconditioners, and ill-conditioning of the overlap matrix. BLOPEX supports parallel MPI-based computations. BLOPEX is incorporated in the HYPRE software and is available as an external package to the PETSc framework.

Anasazi — Anasazi [39] is a well engineered package, with several features that boost robustness and efficiency. Anasazi provides a generic interface to a collection of algorithms for solving large-scale eigenvalue problems. The package implements three methods: a version of the block Krylov-Schur algorithm, a variant of the LOBPCG method with orthogonalization to avoid stability issues, and a block GD method for solving standard and generalized real symmetric and Hermitian eigenvalue problems. All methods are implemented in block variants in order to increase robustness for obtaining multiple eigenvalues and to take advantage of the increased data locality in block matrix-vector, pre-conditioning, and BLAS operations. Even if the total number of matrix-vector multiplications increases in block implementations, for appropriate block sizes this effect is usually balanced by better cache performance.
Anasazi is an interoperable software since both the matrix and vectors are defined as opaque objects so that only knowledge of the matrix and vectors via elementary operations is necessary. As such Anasazi implementations are accomplished via the use of interfaces. Current available interfaces include Epetra. Consequently any libraries that understand Epetra matrices and vectors may be used in conjunction with Anasazi. One of the goals of Anasazi is to allow the user the flexibility in specifying the data representation for the matrix and vectors and so leverage pre-existing software investment. Anasazi is part of the Trilinos framework that includes highly optimized linear algebra operations.

DFT codes implementations — Many of the algorithms previously illustrated have been implemented in variants specifically tailored to DFT codes. Below we give some interesting examples which are far from being exhaustive but can give the reader a flavor of the variety with which iterative methods can be used.
The ABINIT code [40] uses a variant of subspace iteration together with an iterative scheme which improves the initial choice of vectors for the subspace through a form of conjugate gradient algorithm. Orthogonality of the resulting vectors is ensured by a Rayleigh-Ritz method applied to all the subspace each time one eigenvector approximant is modified. ABINIT includes also a block version of the algorithm where parallelism is offered across the several vectors of the block as well as a variant minimizing the residual norm of the subspace vectors.
Among the several diagonalization methods it offers, the Vienna Ab-initio Simulation Package (VASP) [12], uses a Davidson approach with a form of block preconditioning. This method is recommended as a robust alternative to the other methods, though it is also mentioned as being

more costly in some cases. The PARSEC code [41] also uses a modified version of the block Davidson code where the correction equation is preconditioned with a Chebyshev polynomial filter so as to enhance components of the new vectors which will be used to increase the subspace. This method also offers a double restart procedure with an outer and an inner restart loop. The outer restart loop is very similar to a standard implicit restart in keeping the best approximants among the computed subspace. The inner restart loop instead allows the addition of a new block of approximate vectors as soon as some of the sought after eigenpairs have converged. This version of the Davidson method succeeds in better deflating converged vectors and has the added ability to accept approximate solutions in place of the standard augmentation vectors.

Many DFT codes offer some sort of generalization or modified version of a direct inversion of the subspace iteration (DIIS), a method firstly proposed by Pulay [42, 43] in the early '80s. DIIS is a form of Krylov subspace method where an initial subspace is improved through an iterative scheme which individually minimize the residuals of the eigenvectors approximants. In this approach, there is no need to orthogonalize each vector against all others after each update to the basis vectors.

Overall it has been observed by many that Davidson-based algorithms are more robust than methods based on local optimization (like DIIS or Conjugate Gradient). This observation is not a unanimous viewpoint. For example, developers of PWscf and VASP seem to recommend direct minimization, in spite of its less favorable speed. In the end implementations of each specific algorithm is a key factor. With proper implementation, a Davidson- or Krylov-based approach should be vastly superior to direct minimization.

5 DFT-tailored algorithms: an example

As mentioned in the Introduction section, DFT-based methods lead to the self-consistent solution of linearized eigenvalue problems. In other words the non-linear eigenvalue problem generated by the KS equation is solved by a sequence of eigenproblems whose solution is increasingly closer to the one of the original non-linear problem. In practice one starts with a GHEVP $P^{(1)} : A^{(1)}x = \lambda B^{(1)}x$, solves for it, use the solutions to generate the new $P^{(2)}$ so on and so forth. In the end the computational scientist needs to solve a sequence of eigenproblems $\{P^{(1)}, \ldots, P^{(\ell)}, \ldots, P^{(N)}\}$.

It is reasonable to assume — and it has been shown numerically in [44] — that the problems in the sequence are correlated to each other. Thus a particularly tailored eigensolver could take advantage of the correlation to improve performance and scalability. In this section we show an example of such approach specifically designed for DFT methods based on the LAPW basis set. The algorithm choice for this specific set-up is rather peculiar: an iterative eigensolver is selected to solve dense eigenproblems, a pairing which is, in principle, unfavorable. We report below its properties and the numerical performance it achieves.

Chebyshev Filtered Subspace Iteration — Subspace Iteration complemented with a Chebyshev polynomial filter is a well known algorithm in the literature [45]. A version of it was recently developed for a real space dicretization of DFT by Chelikowsky et Al. [46, 47] and included in the PARSEC code. By using a polynomial filter on the initial block of inputed vectors the method experiences a high rate of acceleration. Since the block of vectors spanning

the invariant subspace could easily become linearly dependent the subspace iteration is usually complemented with some re-orthogonalization procedure.
The Chebyshev Filtered Subspace Iteration (ChFSI) algorithm described here is a slightly more sophisticated version of the basic accelerated subspace iteration. This variant is specifically tailored for DFT-like sequences of eigenproblems and it has been developed with the LAPW discretization in mind. In particular the ChFSI algorithm takes advantage of the eigenvectors of the eigenproblem of the (ℓ)-SCF cycle and uses them as input to solve for the eigenproblem at the $(\ell+1)$-SCF cycle.
The whole algorithm is illustrated in the Algorithm 3 scheme. Notice that the initial input is not the initial GHEVP $A^{(\ell)}x = \lambda B^{(\ell)}x$ but its reduction to standard form $H^{(\ell)} = L^{-1}A^{(\ell)}L^{-\mathrm{T}}$ where $B^{(\ell)} = LL^{\mathrm{T}}$, and $\hat{Y}^{(\ell-1)}$ are the eigenvectors of $H^{(\ell-1)}$. ChFSI uses few Lanczos iterations (`line 1`) so as to estimate the upper limit of the eigenproblem spectrum [48]. This estimate is necessary for the correct usage of the filter based on Chebyshev polynomials [45]. After the Chebyshev filter step (`line 3`) the resulting block of vectors is re-orthonormalized using a simple QR factorization (`line 4`) followed by a Rayleigh-Ritz procedure (`line 5`). At the end of the Rayleigh-Ritz step eigenvector residuals are computed, converged eigenpairs are deflated and locked (`line 13`) while the non-converged vectors are sent again to the filter to repeat the whole procedure.

Algorithm 3 Chebyshev Filtered Subspace Iteration with locking

Require: Matrix $H^{(\ell)}$ of the DGEVP reduced to standard form, approximate eigenvectors $\hat{Y}^{(\ell-1)} \doteq \left[\hat{y}_1^{(\ell-1)}, \ldots, \hat{y}_{\text{NEV}}^{(\ell-1)}\right]$ and eigenvalues $\lambda_1^{(\ell-1)}$ and $\lambda_{\text{NEV}+1}^{(\ell-1)}$.
Ensure: Wanted eigenpairs (Λ, Y) .

1: Estimate the largest eigenvalue. ▷ LANCZOS
2: **repeat**
3: Filter the vectors, $\hat{Y} = C_m(\hat{Y})$. ▷ CHEBYSHEV FILTER
4: Re-orthonormalize $\hat{Y}$. ▷ QR ALGORITHM
5: Compute Rayleigh quotient $G = \hat{Y}^{\dagger}H^{(\ell)}\hat{Y}$. ▷ RAYLEIGH-RITZ (Start)
6: Solve the reduced problem $G\hat{W} = \hat{W}\hat{\Lambda}$.
7: Compute $\hat{Y} = \hat{Y}\hat{W}$. ▷ RAYLEIGH-RITZ (End)
8: **for** $i =$ converged $\rightarrow$ NEV **do** ▷ DEFLATION & LOCKING (Start)
9: **if** $r(\hat{Y}_{:,i}, \hat{\Lambda}_i) <$ TOL **then**
10: $\Lambda = \left[\Lambda\ \hat{\Lambda}_i\right]$
11: $Y = \left[Y\ \hat{Y}_{:,i}\right]$
12: **end if**
13: **end for** ▷ DEFLATION & LOCKING (End)
14: **until** converged $\geq$ NEV

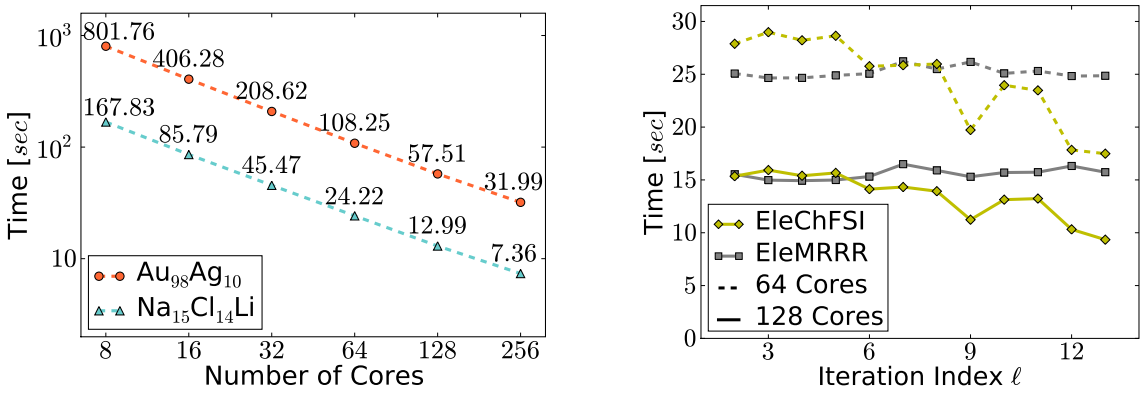
The Chebyshev polynomial filter is at the core of the algorithm. The vectors $\hat{Y}$ are filtered exploiting the 3-terms recurrence relation which defines Chebyshev polynomials of the first kind

$$C_{m+1}(\hat{Y}) = 2\ H\ C_m(\hat{Y}) - C_{m-1}(\hat{Y}) \quad ; \quad C_m(\hat{Y}) \doteq C_m(H)\cdot\hat{Y}. \tag{4}$$

This construction implies all operations internal to the filter are executed through the use of ZGEMM, the most performant among BLAS 3 routines. Since roughly 90% of the total CPU

time is spent in the filter, the massive use of ZGEMM makes ChFSI quite an efficient algorithm and potentially a very scalable one. The parallel MPI version of ChFSI (EleChFSI) is implemented within the Elemental library. The reduced eigenproblem in the Rayleigh-Ritz step is solved using a parallel implementation of the MRRR eigensolver (EleMRRR).

Performance and scalability — In the plots below we report on the scalability of EleChFSI and its performance when compared with the fastest direct method available on the market as shown in [49].
Plot (a) of Fig. 3 illustrate the strong scalability of EleChFSI showing a steady decrease of CPU time as the number of cores increases. The rate of reduction is practically the same for both atomic systems despite their size differ by more than 30%. This plot shows that EleChFSI is extremely efficient even when the ratio of data per processor is not optimal. This result is due both on the re-use of eigenvectors of the previous SCF and on the extensive use of BLAS within the Elemental framework.



(a) EleChFSI Strong scalability.

(b) EleChFSI performance.

Fig. 3: *EleChFSI strong scalability and performance. In plot (a) the size of the eigenproblems are kept fixed while the number of cores is progressively increased. Eigenproblems of size* $n = 13,379$ *and* $n = 9,273$ *are shown. In plot (b) EleChFSI is compared with EleMRRR on eigenproblems of increasing self-consistent cycle index* ℓ *for a sequence of eigenproblems with* $n = 9,273$.

Compared to direct solvers, EleChFSI promises to be quite competitive. Depending on the number of eigenpairs computed, the algorithm implementation is even faster than EleMRRR. In plot (b) of Fig. 3 EleChFSI is already faster than EleMRRR for half of the eigenproblems in the sequence (64 cores). When the tests are repeated with 128 cores EleChFSI is inequivocably the faster of the two algorithms. Since the fraction of the spectrum computed in plot (b) is $\sim 3\%$, Fig. 3 shows that EleChFSI scales better than EleMRRR and is more performant when the sought number of eigenpairs is not too high.

References

[1] P. Hohenberg, Physical Review **136**(3B), B864 (1964).

[2] W. Kohn and L. J. Sham, Phys.Rev. **140**, A1133 (1965).

[3] S. Huzinaga, The Journal of Chemical Physics **42**(4), 1293 (1965).

[4] J. A. Pople and W. J. Hehre, Journal of Computational Physics **27**(2), 161 (1978).

[5] E. Wimmer, M. Weinert, and A. J. Freeman, Physical Review B **24**(2), 864 (1981).

[6] H. J. F. Jansen and A. J. Freeman, Physical Review B **30**(2), 561 (1984).

[7] J. Chelikowsky, N. Troullier, K. Wu, and Y. Saad, Physical Review B **50**(16), 11355 (1994).

[8] S. White, J. Wilkins, and M. Teter, Physical Review B **39**(9), 5819 (1989).

[9] L. Genovese, A. Neelov, S. Goedecker, T. Deutsch, S. A. Ghasemi, A. Willand, D. Caliste, O. Zilberberg, M. Rayson, and A. Bergman, The Journal of Chemical Physics **129**, 014109 (2008).

[10] L. S. Blackford, J. Choi, A. Cleary, E. D'Azevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, *et al.*, *ScaLAPACK Users' Guide* (Society for Industrial and Applied Mathematics, 1987).

[11] R. Lehoucq, D. Sorensen, and C. C. Yang, *ARPACK Users' Guide: Solution of Large-scale Eigenvalue Problems* (Society for Industrial and Applied Mathematics, 1998).

[12] G. Kresse and J. Furthmüller, Physical Review B **54**(16), 11169 (1996).

[13] Y. Saad, Y. Zhou, C. Bekas, M. L. Tiago, and J. Chelikowsky, physica status solidi (b) **243**(9), 2188 (2006).

[14] G. M. Amdahl, in *the April 18-20, 1967, spring joint computer conference* (ACM Press, New York, New York, USA, 1967), p. 483.

[15] J. L. Gustafson, Communications of the ACM **31**(5), 532 (1988).

[16] G. H. Golub and C. F. Van Loan, *Matrix Computations* (Johns Hopkins Univ., 2012).

[17] C. H. Bischof, B. Lang, and X. Sun, ACM Transactions on Mathematical Software **26**(4), 581 (2000).

[18] D. S. Dodson and J. G. Lewis, ACM Signum Newsletter **20**(1), 19 (1985).

[19] J. J. Dongarra, J. Du Croz, S. Hammarling, and R. J. Hanson, ACM Transactions on Mathematical Software **14**(1), 18 (1988).

[20] J. J. Dongarra, J. Du Croz, S. Hammarling, and I. S. Duff, ACM Transactions on Mathematical Software **16**(1), 1 (1990).

[21] K. Goto and R. A. v. d. Geijn, ACM Transactions on Mathematical Software **34**(3), 1 (2008).

[22] K. Goto and R. van de Geijn, ACM Transactions on Mathematical Software **35**(1), 1 (2008).

[23] R. C. Whaley and J. J. Dongarra, *Automatically tuned linear algebra software* (IEEE Computer Society, 1998).

[24] J. W. Demmel, O. A. Marques, B. N. Parlett, and C. Vömel, SIAM Journal on Scientific Computing **30**(3), 1508 (2008).

[25] M. Petschow, E. Peise, and P. Bientinesi, SIAM Journal on Scientific Computing **35**(1), C1 (2013).

[26] T. Auckenthaler, H. J. Bungartz, T. Huckle, L. Krämer, B. Lang, and P. Willems, Journal of Computational Science **2**(3), 272 (2011).

[27] J. Poulson, B. Marker, R. A. van de Geijn, J. R. Hammond, and N. A. Romero, ACM Transactions on Mathematical Software (TOMS) **39**(2) (2013).

[28] P. Bientinesi, I. S. Dhillon, and R. A. van de Geijn, SIAM Journal on Scientific Computing **27**(1), 43 (2005).

[29] D. S. Watkins, *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods* (Society for Industrial and Applied Mathematics, 2007).

[30] L. Giraud, J. Langou, M. Rozložník, and J. v. d. Eshof, Numerische Mathematik **101**(1), 87 (2005).

[31] A. V. Knyazev, SIAM Journal on Scientific Computing **23**(2), 517 (2001).

[32] A. V. Knyazev, M. E. Argentati, I. Lashuk, and E. E. Ovtchinnikov, SIAM Journal on Scientific Computing **29**(5), 2224 (2007).

[33] E. Di Napoli and M. Berljafa, Computer Physics Communications **184**(11), 2478 (2013).

[34] V. Hernandez, J. E. Roman, and V. Vidal, ACM Transactions on Mathematical Software **31**(3), 351 (2005).

[35] G. W. Stewart, SIAM Journal on Matrix Analysis and Applications **23**(3), 601 (2002).

[36] G. W. Stewart, SIAM Journal on Matrix Analysis and Applications **24**(2), 599 (2002).

[37] S. Balay, W. Gropp, L. Curfman McInnes, and B. Smith, *PETSc 2.0 users manual: Revision 2.0.16* (Argonne National Laboratory, 1997).

[38] A. Stathopoulos and J. R. McCombs, ACM Transactions on Mathematical Software (TOMS) **37**(2), 21 (2010).

[39] C. G. Baker, U. L. Hetmaniuk, R. B. Lehoucq, and H. K. Thornquist, ACM Transactions on Mathematical Software (TOMS) **36**(3), 13 (2009).

[40] X. Gonze, B. Amadon, P. M. Anglade, J. M. Beuken, F. Bottin, P. Boulanger, F. Bruneval, D. Caliste, R. Caracas, M. Côté, T. Deutsch, L. Genovese, *et al.*, Computer Physics Communications **180**(12), 2582 (2009).

[41] L. Kronik, A. Makmal, M. L. Tiago, M. M. G. Alemany, M. Jain, X. Huang, Y. Saad, and J. R. Chelikowsky, physica status solidi (b) **243**(5), 1063 (2006).

[42] P. Pulay, Chemical Physics Letters **73**(2), 393 (1980).

[43] P. Pulay, Journal of Computational Chemistry **3**(4), 556 (1982).

[44] E. Di Napoli, S. Blügel, and P. Bientinesi, Computer Physics Communications **183**(8), 1674 (2012).

[45] Y. Saad, *Numerical methods for large eigenvalue problems* (Siam, 2011).

[46] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, Journal of Computational Physics **219**(1), 172 (2006).

[47] Y. Zhou, Y. Saad, M. L. Tiago, and J. R. Chelikowsky, Physical Review E **74**(6), 066704 (2006).

[48] Y. Zhou and R.-C. Li, Linear Algebra and its Applications **435**(3), 480 (2011).

[49] M. Berljafa and E. Di Napoli, arXiv.org (2013), `1305.5120v1`.

D 4 DFT in practice [1]

D. Wortmann
Institute for Advanced Simulation and
Peter Grünberg Institut
Forschungszentrum Jülich GmbH

Contents

[1] Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercom-

1 Introduction

The evolution of density functional theory (DFT) into the standard model of many fields of modern physics, chemistry and material sciences would not have been possible without the invention of numerical algorithms and their implementation into powerful computer codes. The development of these methods and codes is driven by an ever increasing demand to simulate larger systems with more complex properties using increasingly sophisticated approximations as discussed in many talks of this spring school. On the other hand, the huge progress in computing hardware and especially the supercomputing development – being in a state of permanent revolution – imposes huge challenges for application development. Most challenging is the conflict between an old code base and modern concepts of programming: on the one hand modern DFT codes tend to include an enormous wealth of functionality which needs a careful and often painfully slow adaptation to new computational paradigms, on the other hand these paradigms tend to be in constant flow.

The quest to solve the Kohn-Sham equation (6) efficiently for periodic solids, solids with surfaces and interfaces, clusters and molecules has led to a wide spectrum of very successful and efficient electronic structure methods. Treating isolated clusters or molecules, methods based on localized orbitals are frequently selected going hand in hand with the chemical intuition of a system in question. Considering methods applicable to periodic solids, frequently algorithms are chosen where the Bloch boundary condition can be included in the basis set. Guiding principles to develop electronic structure methods are obtained by having a closer look at the mathematical nature of the Schrödinger-like Kohn-Sham equation Eq. (6). Here the numerical most demanding contributions arise from the kinetic energy operator Δ and the $1/r$ singularity at the nucleus with the simultaneous necessity to calculate the xc-potential $V_{\mathrm{xc}}[n](\mathbf{r})$ and the Hartree potential $V_{\mathrm{H}}[n](\mathbf{r})$.

Within this contribution to the spring school I will concentrate on numerical and algorithmic aspects of codes implementing DFT. In this sense, this manuscripts complements the one of M. Ležaić (A2) which provides a general introduction into the formalism. Furthermore, only the most basic problems one has to solve in a standard DFT code are discussed. Many of the more advanced schemes also discussed within other contributions of the school, which are based on DFT or can be considered as extensions of DFT, have their own set of numerical challenges and often the general considerations discussed here have to be extended or modified accordingly. In particular, while the solution of the Kohn-Sham eigenvalue equation is the key task in most DFT codes, in terms of computational time needed this part can become less relevant in methods in which very time consuming schemes beyond standard DFT are employed. These much more demanding methods include all kinds of hybrid functional approaches or the GW approximation.

Hence, within this manuscript four main points will be discussed: (i) while giving a brief primer into DFT many smaller algorithmic points are mentioned, (ii) the problem of the all-electron potential, the core states and the pseudo-potential approach are described, (iii) different basis sets and the Green function method to solve the Kohn-Sham equation are mentioned and (iv) a concluding overview of different codes is presented and the computationally expensive parts are identified.

Within the manuscript atomic Hartree units with $\hbar = m = 1$ are used.

2 Brief primer of Kohn-Sham DFT

As the fundamentals of density functional theory was already introduced in detail by M. Ležaić (A2) only the most basic equations and procedures are sketched here for completeness with a focus on the computational aspects of the DFT self-consistency cycle.

Total Energy and Force

In the density-functional theory, the total energy $E[\{\mathbf{R}_j\},\{\psi_\nu\}]$ of a system of interacting atoms and electrons is a functional of the atomic positions $\{\mathbf{R}_j\}$ and the electron density $n(\mathbf{r})$. The electron density can be expressed in terms of M occupied single-particle orbitals $\psi_\nu(\mathbf{r})$:

$$n(\mathbf{r}) = \sum_{\nu(occ)}^{M} \mid \psi_\nu(\mathbf{r}) \mid^2, \tag{1}$$

where ν labels the states. If the total energy functional $E[\{R\},\{\psi_\nu\}]$ is minimized with respect to the electronic degrees of freedom $\{\psi_i\}$, we recover the Born-Oppenheimer surface $\Phi[\{R\}]$:

$$\Phi[\{\mathbf{R}_j\}] = \min_{\{\psi_\nu\}} E[\{\mathbf{R}_j\},\{\psi_\nu\}], \tag{2}$$

on which the atoms move. The derivative of $\Phi[\{\mathbf{R}_j\}]$ with respect to the atomic position $\mathbf{R}_j$ gives the force $\mathbf{F}_j$,

$$\mathbf{F}_j = -\nabla_{\mathbf{R}_j}\Phi[\{\mathbf{R}_j\}] \tag{3}$$

exerted on the atom j, which ties electronic structure to structural optimization and molecular dynamics calculations.

The Kohn-Sham Equations

The single-particle wavefunctions $\psi_\nu(\mathbf{r})$ are obtained by minimization of the total energy with respect to the wavefunctions subject to the normalization constraint

$$\int dr \mid \psi_\nu(\mathbf{r}) \mid^2 = 1. \tag{4}$$

This leads to the Kohn-Sham equations[1], an eigenvalue problem for the eigenfunctions $\psi_i(r)$ and the eigenvalues ε_i:

$$\hat{H}[n]\,\psi_\nu[n] = \varepsilon_\nu[n]\,\psi_\nu[n], \tag{5}$$

where all quantities depend on the electron density n. The Hamiltonian $\hat{H}$ is a sum of three main terms and the eigenvalue problem is written in the form:

$$(\hat{T}_0 + \hat{V}_{\text{ext}} + \hat{V}_{eff})\,\psi_\nu(\mathbf{r}) = \varepsilon_\nu\,\psi_\nu(\mathbf{r}) \tag{6}$$

In the real space representation the individual terms are the following:

$$\text{kinetic energy :} \qquad \hat{T}_0 = -\frac{\hbar^2}{2m}\Delta_{\mathbf{r}} \tag{7}$$

$$\text{external-potential :} \qquad V_{\text{ext}}(\{\mathbf{R}_j\},\mathbf{r}) = \sum_\mu \frac{e^2 Z_j}{|\mathbf{r}-\mathbf{R}_j\mid} \tag{8}$$

$$\text{effective potential :} \qquad V_{eff}(\mathbf{r}) = V_{eff}[n](\mathbf{r}) \tag{9}$$

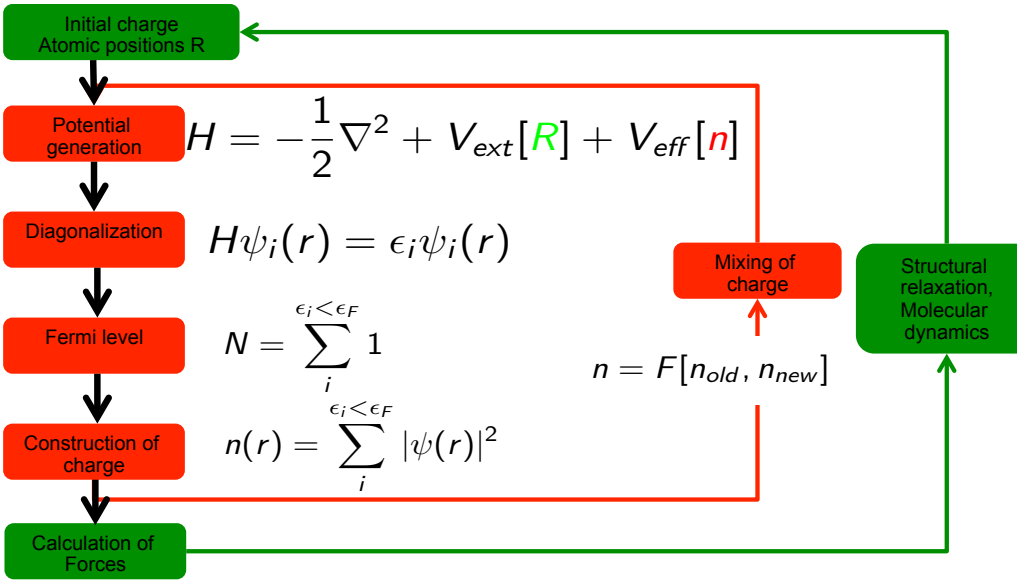


Fig. 1: *Typical flow of a DFT calculation.*

The effective potential in this equation describes the interaction of the electrons and is an potential depending explicitly on the density. Thus, the Hamiltonian $\hat{H}[n]$ and the wavefunctions $\psi_\nu([n], \mathbf{r})$ are also dependent on the electron density $n(r)$. Together with the expression Eq. (1) a self-consistency problem to obtain the charge density $n(\mathbf{r})$ is established, which is solved iteratively until the input density (used to define the potential terms in the Hamiltonian) is equal to the output density within the required accuracy.
The external potential $\hat{V}_{\text{ext}}[\{\mathbf{R}\}]$ depends explicitly on the positions $\{\mathbf{R}\}$ of all atoms, which change at certain steps to optimize the atomic structure or every time-step of a molecular dynamics algorithm. Thus, the Hamiltonian $\hat{H}[\{R\}]$ and the wavefunctions $\psi_\nu(\{\mathbf{R}\}, \mathbf{r})$ are also dependent on the atomic positions $\{\mathbf{R}\}$. After the self-consistency condition for the electron density has been fulfilled, the atom positions are moved by a molecular static or molecular dynamics time-step, $\{R(t)\} \rightarrow \{R(t + \Delta t)\}$. Thus, for N_{MD} molecular time steps the eigenvalue problem has to be solved $N_{\text{MD}} N_{\text{iter}}$ times. These arguments suggest a particular loop structure of a typical first-principles method and a particular sequence how the different elements are calculated. This is summarized in Fig. 1.

The Effective Single Particle Potential

Of crucial importance and key to the success of DFT is the effective potential describing the interaction of the many-electron system. It is constructed as a sum of two ingredients, (i) the Hartree potential describing the classical electrostatic interaction of a single electron with the

charge of the system and (ii) the exchange-correlation potential which can be simply viewed as the remaining term: $V_{eff}[n] = V_H[n] + V_{xc}[n]$.
The main advantage of this representation is the possibility to calculate the Hartree potential by solving Poisson's equation$\Delta_r V_H[n](\mathbf{r}) = 4\pi e^2 n(\mathbf{r})$. While the solution of Poisson's equation is definitely not trivial, this is probably one of the best studied problems in computational physics and fast and efficient methods such as Fast-Fourier Transforms (FFTs) or multi-grid methods are employed to solve this equation.
For the remaining exchange-correlation potential many different approximations are available of which the Local Density Approximation (LDA) and Generalized Gradient Approximations (GGA) are simplest and most widely used due to their astonishing accuracy and their extreme simplicity. In these approximations, simple formulas are available to evaluate the potential at each point in space by the density (and derivative of the density) at the very same point. Thus, these approximations generate computationally very cheap algorithms. As introduced in several contributions to this school, a whole zoo of approximations beyond these established DFT approximations have been invented which usually can be viewed as a replacement of the simple LDA or GGA exchange-correlation potential by a more sophisticated expression. Frequently, these advanced approaches do not lead to a simple local potential but to an non-local, often orbital or energy dependent quantity which not only is more involved to calculate but also might modify the mathematical structure of the Kohn-Sham equation itself. These approaches then generally require modifications of the procedures discussed in this manuscript and tend to introduce computationally expensive algorithms different from standard DFT calculations which I will not discuss here further.
If magnetism occurs, the ground state has a broken symmetry and the ground-state energy is described by functionals which depend on the (vector-)magnetization density $\mathbf{m}(\mathbf{r})$ as an additional field to the ordinary charge density $n(r)$. An additional term $\mu_B \underline{\sigma} \cdot \mathbf{B}_{\mathrm{xc}}(\mathbf{r})$ appears in the Kohn-Sham equations Eq. (6), where $\mu_{\mathrm{B}} = \frac{e\hbar}{2mc}$ is the Bohr magneton, $\mathbf{B}_{\mathrm{xc}}$ is the magnetic xc-field an electron experiences, and $\underline{\sigma}$ are the Pauli spinors. Thus, calculating magnetic systems, one works in a two-dimensional spin-space. The Hamiltonian is a 2×2 matrix in spin-space. Complex magnetic structures lower frequently the symmetry of the problem and more states have to be calculated, pushing the computational effort to the limits of modern supercomputers. In case of collinear magnetism, e.g. ferro-, ferri-, or anti-ferromagnetism (and without considering relativistic effects), the Hamiltonian is diagonal in spin space, the magnetization density m_z is then given by spin-up and -down densities, $m_z(\mathbf{r}) = n_\uparrow(\mathbf{r}) - n_\downarrow(\mathbf{r})$, and the effort of a magnetic calculation is just twice that of a nonmagnetic one. In general, the vector magnetization density can not be written as $\mathbf{m}(r) = m_z(\mathbf{r})\mathbf{e}_z$ with $\mathbf{e}_z$ being the unit-vector defining the collinear spin-quantization axis, and the search of the magnetic structure can be done dynamically bearing similarities to the dynamical structure optimization combining molecular dynamics and simulated annealing. Throughout this manuscript, the spin label is dropped for convenience. More information on the treatment of magnetism can be found in the chapter C4 "Complex Magnetism" by S. Blügel.

2.1 The Eigenvalue Problem

One of the key parts of any DFT code is of course the computational solution of the Kohn-Sham equation itself. While there are many possible ways to solve the Kohn-Sham equations the probably most frequently employed scheme uses a variational method in which a basis set is introduced. The set of wavefunctions $\psi_\nu(r)$ with the index ν is sought as a linear combination

of basis functions $\varphi_n(r)$

$$\psi_\nu(\mathbf{r}) = \sum_{n=1}^{N} c_\nu^n \varphi_n(\mathbf{r}) \tag{10}$$

where the c_ν^n are the expansion coefficients of the wavefunction (coefficient vector), and N is the number of basis functions taken into account. Using this expansion, the eigenvalue problem

$$\hat{H}\psi_\nu(\mathbf{r}) = \varepsilon_\nu \psi_\nu(\mathbf{r}) \tag{11}$$

is translated into an algebraic generalized eigenvalue problem of dimension N

$$\begin{pmatrix} H^{1,1} & H^{1,2} & \dots & H^{1,N} \\ H^{2,1} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ H^{N,1} & \dots & \dots & H^{N,N} \end{pmatrix} \begin{pmatrix} c_\nu^1 \\ c_\nu^2 \\ \vdots \\ c_\nu^N \end{pmatrix} = \epsilon_\nu \begin{pmatrix} S^{1,1} & S^{1,2} & \dots & S^{1,N} \\ S^{2,1} & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \vdots \\ S^{N,1} & \dots & \dots & S^{N,N} \end{pmatrix} \begin{pmatrix} c_\nu^1 \\ c_\nu^2 \\ \vdots \\ c_\nu^N \end{pmatrix} \tag{12}$$

for the coefficient vector c_ν^n corresponding to the eigenvalues ε_ν. The Hamilton matrix $H = \{H^{n,n'}\}$ and overlap matrix $S = \{S^{n,n'}\}$ are in general hermitian (For special symmetry and/or boundary conditions of the problem they might also be chosen to be real symmetric). The Hamiltonian matrix (or short Hamiltonian) is given in terms of its matrix elements defined as the integral over the cell-volume Ω

$$H^{n,n'} = \int_\Omega \varphi_n^*(\mathbf{r}) \hat{H}(\mathbf{r}) \varphi_{n'}(\mathbf{r}) d^3r \tag{13}$$

If the basis functions are orthonormal, i.e. $\langle \varphi_n | \varphi_{n'} \rangle = \delta^{n,n'}$, then the overlap matrix with its matrix elements

$$S^{n,n'} = \int_\Omega \varphi_n^*(\mathbf{r}) \varphi_{n'}(\mathbf{r}) d^3r \tag{14}$$

becomes diagonal, $S^{n,n'} = \delta^{n,n'}$, and the generalized eigenvalue problem Eq. (12) becomes of standard type.

In general, the generalized eigenvalue problem is reduced to a standard one using the Cholesky decomposition. It can be shown (e.g. Stoer [2]), that any hermitian and positive definite matrix can be decomposed into a matrix product of a lower triangular with only positive diagonal elements matrix and its transposed. Clearly, the overlap matrix satisfies these conditions and can be written $S = LL^H$. Therefore, Eq. (12) becomes

$$Hc_\nu = \varepsilon_\nu L L^H c_\nu, \tag{15}$$

multiplying from the left with L^{-1} and introducing a unit matrix $I = (L^{-1})^H L^H$ we obtain

$$Px_\nu = \varepsilon_\nu x_\nu, \tag{16}$$

after we have defined $P = L^{-1}H(L^{-1})^H$ and $x_\nu = L^H c_\nu$. Thus, the generalized eigenvalue problem has been reduced to a simple one. The eigenvectors c_ν can be obtained by the back-transformation, $c_\nu = (L^H)^{-1} x_\nu$.

The choice of the most efficient numerical algorithm to solve Eq. (12) depends on the number of basis functions N and the number M of states ν taken into account. If $M/N >\sim 0.1$, direct numerical diagonalization schemes are usually employed. If $M/N <\sim 0.1$ or if N is too large to fit the eigenvalue problem into the memory of a computer the eigenvalue problem can be solved iteratively. For more details of these different strategies in solving the eigenvalue problem and their parallelisation the reader is referred to the contribution of E. Di Napoli (D3).

2.2 Brillouin-Zone Integration and Fermi Energy

The calculation of quantities like the electron density, total energy or forces requires to sum over all solutions of the eigenvalue problem, so far labeled with the index ν. These sums are central to the DFT method. As already mentioned, for collinear magnetic calculations, the solutions can be classified by their spin, so that the sums contain a spin summation. Additionally, in periodic structures one exploits the Bloch theorem so that in general $\nu = (\sigma, \mathbf{k}, i)$ should be viewed as a composite index containing the spin, the $\mathbf{k}$-value and a band index. The $\mathbf{k}$-vector is special as it corresponds to a continuos variable which has to be integrated over the Brillouin zone (BZ).

$$\sum_{\nu} \longrightarrow \frac{1}{V_{BZ}} \sum_{\sigma=\uparrow,\downarrow} \int_{BZ} d^3k \sum_{i}$$

These sums/integrations stretch only over the occupied states, i.e. over those states with energy $\epsilon_\nu = \epsilon(\sigma, \mathbf{k}, \mathrm{i})$ lower than the Fermi energy E_F. The Fermi energy can be determined from the requirement that the total number of electrons N of the system is given by

$$N = \frac{1}{V_{BZ}} \sum_{\sigma=\uparrow,\downarrow} \int_{BZ} d^3k \sum_{i}^{\epsilon(\sigma,\mathbf{k},i)<E_F} 1. \tag{17}$$

Numerically, the integration over the $\mathbf{k}$-vector is performed on a discrete mesh in the Brillouin zone. In fact the effort of the BZ integration is in practice significantly reduced by employing the point group symmetry, where the integration is reduced to the irreducible wedge of the BZ (IBZ). There are different methods, that can be used to perform the integration, e.g. the special points method[3, 4] and the tetrahedron method[5, 6, 7]. The special points method is a method to integrate smoothly varying periodic functions of $\mathbf{k}$. The function to be integrated has to be calculated at a set of special points in the IBZ, each of which is assigned a weight. Thus, the BZ integration is transformed into a sum over a set of $\mathbf{k}$-points. Thus, the integrals become:

$$\frac{1}{V_{BZ}} \int_{BZ} d^3k \longrightarrow \sum_{\mathbf{k}\in \mathrm{IBZ}} w(\mathbf{k}) \tag{18}$$

The special k-points method does not converge very quickly, and rather many $\mathbf{k}$-points are needed to obtain accurate results. On the other hand this method is simple to implement, because the weights depend only on each $\mathbf{k}$-point. Frequently, numerical problems arise from the "sharp" differentiation between occupied and empty bands (parts of bands). Let's consider a band that is very close to the Fermi energy at a certain $\mathbf{k}$-point. During the iterations the energy of this band might rise above or drop below the Fermi energy. This leads to sudden changes in the charge density, which can slow down or even prevent the convergence of the density. These sudden changes are clearly a result of the discretization in momentum space. To avoid this problem, the sharp step at the Fermi energy can be smoothened, e.g. by introducing a so-called temperature broadening in the context of a the Fermi function $(e^{(\epsilon-E_F)/k_BT}+1)^{-1}$ so that finally the sum over all states turns into

$$\sum_{\nu} \longrightarrow \sum_{\sigma=\uparrow,\downarrow} \sum_{\mathbf{k}\in \mathrm{IBZ}} w(\mathbf{k}) \sum_{i}^{\infty} (e^{(\epsilon(\sigma,\mathbf{k},i)-E_F)/k_BT}+1)^{-1}$$

where the temperature T is an additional external parameter adjusted to obtain the good convergence without smoothing too much.

2.3 Achieving Self-Consistency

As last point in our short general overview of DFT I should shortly discuss the problem of obtaining self-consistency. By this term the condition is understood that the density $n_{in}(\mathbf{r})$ used to construct the effective potential $V_{\text{eff}}(\mathbf{r})[n_{in}]$ and the density field $n_{out}(\mathbf{r}) = \sum_\nu |\psi_\nu(\mathbf{r})|^2$ are consistent. In other words, the electron density $n_0(\mathbf{r})$ that minimizes the energy functional is a fix-point of the mapping

$$n_{out}(\mathbf{r}) = F\{n_{in}(\mathbf{r})\}. \tag{19}$$

i.e. it solves

$$\mathcal{F}\{n_0(\mathbf{r})\} = 0, \quad \text{with } \mathcal{F}\{n(\mathbf{r})\} = F\{n(\mathbf{r})\} - n(\mathbf{r})\}. \tag{20}$$

(The same can be formulated for the potential.) Typically, the density is expanded into a large set of basis functions. Thus, in actual calculations, the charge density is a coefficient vector $\mathbf{n}$ of dimension $N_{\text{Q}} \sim 8 * N$ (N being again the number of basis functions as defined in Eq. (10)) and Eq. (20) constitutes a system of N_{Q} nonlinear equations, which can be solved by iteration:

$$\mathbf{n}^{(m+1)} = F\{\mathbf{n}^{(m)}\}. \tag{21}$$

A usually very good starting density $\mathbf{n}^{(0)}$ can be constructed by a superposition of atomic densities. A straight mapping as is suggested in Eq. (21) however is in general divergent and convergence can be achieved only if the output density $F\{\mathbf{n}^{(m)}\}$ is mixed with the input density $\mathbf{n}^{(m)}$.
The simplest and slowest of such mixing schemes is the so-called "simple mixing", which converges only linearly. The density for the next iteration is constructed as a linear combination of $\mathbf{n}^{(m)}$ and $F\{\mathbf{n}^{(m)}\}$ according to:

$$\mathbf{n}^{(m+1)} = (1-\alpha)\mathbf{n}^{(m)} + \alpha F\{\mathbf{n}^{(m)}\} = \mathbf{n}^{(m)} + \alpha\mathcal{F}\{\mathbf{n}^{(m)}\}, \tag{22}$$

where α is the so-called mixing parameter. If it is chosen small enough, the iteration converges and is very stable. However, for the type of systems one is frequently interested in, α tends to be very very small, requiring many hundreds of iterations. In spin-polarized calculations different mixing parameters can be used for the charge and the magnetization density. Usually, the spin mixing parameter can be chosen far larger than the parameter for the charge density.
More efficient than a simple mixing scheme are Newton-Raphson methods, in which the functional $\mathcal{F}\{n\}$ is linearized around the approximate solution $\mathbf{n}^{(m)}$.

$$\mathcal{F}\{\mathbf{n}\} \approx \mathcal{F}\{\mathbf{n}^{(m)}\} + \mathcal{J}\{\mathbf{n}^{(m)}\}(\mathbf{n} - \mathbf{n}^{(m)}), \quad \mathcal{J}\{\mathbf{n}^{(m)}\} = \left.\frac{\partial\mathcal{F}\{\mathbf{n}\}}{\partial\mathbf{n}}\right|_{\mathbf{n}^{(m)}}. \tag{23}$$

In actual calculations the Jacobian $\mathcal{J}$ is a $N_{\text{Q}} \times N_{\text{Q}}$ matrix. Similar to the well-known Newton method to find zeros of one-dimensional functions, the next approximation to $\mathbf{n}_0$, $\mathbf{n}^{(m+1)}$, is determined from the requirement, that the linearized functional in Eq. (23) vanishes at $\mathbf{n}^{(m+1)}$. Thus, $\mathbf{n}^{(m+1)}$ is given by:

$$\mathbf{n}^{(m+1)} = \mathbf{n}^{(m)} - \left[\mathcal{J}\{\mathbf{n}^{(m)}\}\right]^{-1}\mathcal{F}\{\mathbf{n}^{(m)}\}. \tag{24}$$

In opposite to the simple mixing, the Newton-Raphson method converges quadratically. The major drawback of this method is the difficulty to evaluate the Jacobian. Even if the functional $\mathcal{F}\{\mathbf{n}\}$ were known more explicitly, the evaluation would be cumbersome due to the enormous

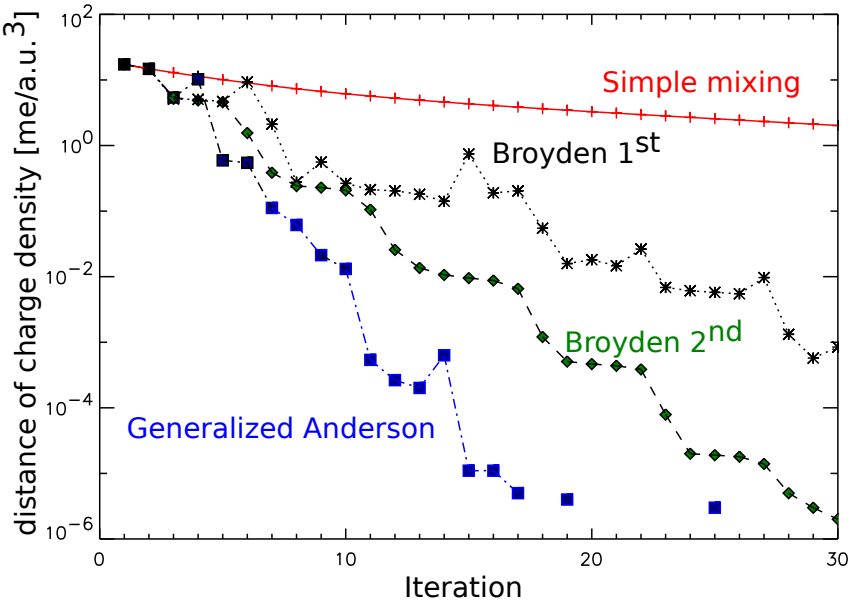


Fig. 2: *Comparison of the convergence of a non-magnetic bcc Fe calculation using the LAPW method. Different mixing schemes are used and the distances are plotted logarithmically[11].*

size of $\mathcal{J}\{\mathbf{n}\}$. In addition, the Jacobian has to be inverted where the amount of calculation scales with cube of the dimension. A further problem is that the convergence radius of this method is rather small so that it can only be used if $\mathbf{n}^{(m)}$ is already very close to $\mathbf{n}_0$.
The development of the Quasi-Newton methods made it possible to exploit the advantages of the Newton-Raphson method, i.e. to make use of the information that is contained in the Jacobian, for problems where the Jacobian cannot be calculated or its determination is too demanding. Rather than computing the Jacobian each iteration, an approximate Jacobian is set up and improved iteration by iteration. From the linearization of $\mathcal{F}\{\mathbf{n}\}$ in Eq. (23) we find the following condition for the Jacobian, which is usually called Quasi-Newton condition:

$$\Delta\mathbf{n}^{(m)} = \left[\mathcal{J}^{(m)}\right]^{-1} \Delta\mathcal{F}^{(m)} \tag{25}$$

$$\Delta\mathbf{n}^{(m)} = \mathbf{n}^{(m)} - \mathbf{n}^{(m-1)}, \quad \Delta\mathcal{F}^{(m)} = \mathcal{F}\{\mathbf{n}^{(m)}\} - \mathcal{F}\{\mathbf{n}^{(m-1)}\}$$

Quasi-Newton methods converge super-linearly and have a larger convergence radius than the Newton-Raphson method. Since the Jacobian is build up iteration by iteration, the "history" of the previous iterations is memorized in $\mathcal{J}$, whereas the Jacobian of the Newton-Raphson method depends only on the previous iteration. In this sense the Newton-Raphson method is self-corrective[8], it "forgets" inadequately chosen corrections. The Quasi-Newton methods sometimes need to be restarted, if the iteration converges only slowly. This can happen if the starting density is very far from $\mathbf{n}_0$ or when physical or numerical parameters that affect the calculations are changed during the iteration. Eq. (25) does not determine the Jacobian uniquely, instead Eq. (25) constitutes a system of N_Q equations for N_Q^2 unknowns. The various Quasi-Newton schemes differ by the ansatz how the new information is used to build the inverse Jacobian. The quality of the convergence is measured by the distance of the residual vector:

$$d_{\mathbf{n}^{(m)}} = ||\mathcal{F}\{\mathbf{n}^{(m)}\}|| = ||F\{\mathbf{n}^{(m)}\} - \mathbf{n}^{(m)}||. \tag{26}$$

The correct choice of a mixing scheme can significantly influence the convergence behavior of the calculation and thus can determine the computational effort needed (cf. Fig. 2). Be-

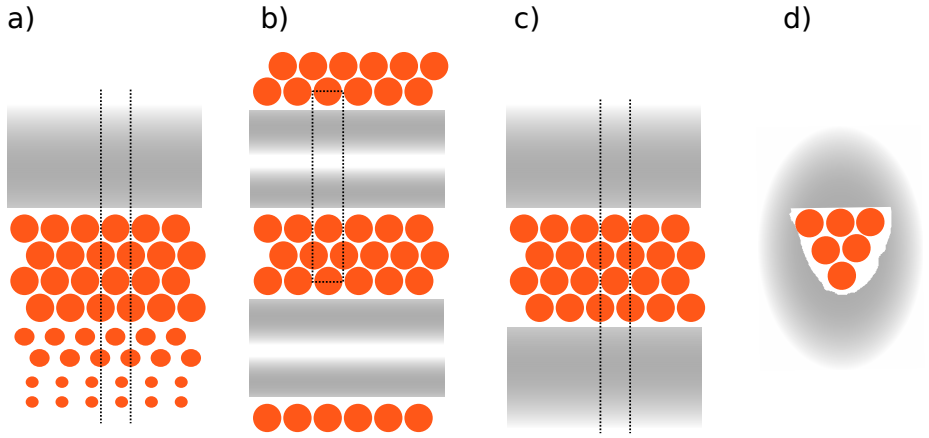


Fig. 3: *Different possible structural models for surface calculations. The vacuum area is depicted by a shaded gray, indicating the decay of the charge. a-c) show a semi-infinite surface model, the repeated slab geometry and the single-slab model, respectively. In these models the unit-cell is indicated by dashed lines. d) shows an isolated cluster model.*

sides accelerating the convergence by the use of generalized Newton methods also schemes for preconditioning are employed[9, 10].

2.4 Structural models: surfaces and defects

One of the most crucial steps in actual DFT calculations is the proper choice of the structural model. As discussed in the introductory lecture (A1) of the school one can efficiently treat insulated systems, i.e. finite systems in which the Kohn-Sham equation is solved in a finite volume, and infinite but periodic systems in which one exploits the Bloch theorem. Of course these cover only a small fraction of possible geometries and many problems of most significance are actually somehow "in between" these two limiting cases. For example, the surface of a crystal on the one hand breaks the periodicity of the system and thus cannot be calculated by constructing Bloch states only, on the other hand a surface can also not be easily considered as an insulated system of finite size. Similar problems arise in the treatment of defects in crystals such as impurities, vacancies, stacking faults, domain boundaries, grain boundaries and many more.

Fortunately, many – but by no means all – phenomena in such structures are relatively short-range in nature. This makes it possible to choose geometric models which are small enough to be tractable with electronic structure methods yet still large enough to be physically meaningful. Hence, one often can construct models which are derived from the basic periodic or insulated geometries containing a limited set of atoms. The construction of such models is of course highly depended on the system and the property one wants to study. In the following, I will discuss the most common geometric models for electronic structure calculations of surfaces and outline their range of applicability.

Conceptually the most satisfying surface geometry is that of a semi-infinite solid in which a periodic infinite crystal is cut at some plane and the second half-space is replaced by (semi-infinite) vacuum (cf. Fig. 3a)). While this geometry can be treated relatively easily if one restricts it to a one-dimensional problem, the use of a semi-infinite solid is much more difficult if a full

three-dimensional solution of the DFT problem is attempted.
However, it is reasonable to assume that any material becomes bulk-like at a certain distance away from the surface. A priori, one does not know that distance, but the electronic screening length is a good measure. This results to about 10 layers underneath the surface for transition metals and semiconductors and about 20 layers for sp-metals such as Al, Bi or Pb. In the top layers or the so-called "surface region", the electronic wave functions are then chosen to match the bulk states inside the solid and satisfy the vacuum boundary conditions above the surface. Green function techniques are used, for example, in the Korringa, Kohn and Rostoker (KKR) or in the Full-Potential Linearized Augmented (FLAPW) Green function methods [12, 13], which provide the necessary mathematical apparatus to accomplish this matching procedure [14].
A simple, but effective geometric surface model is the repeated slab geometry (cf. Fig. 3b)) calculations of surfaces. Thin films consisting of about 10 to 20 layers are repeated in the direction perpendicular to the surface. The slabs are chosen thick enough to approach bulk-like behavior near the center of each slab and the spacing is taken large enough so that any artificial interactions across the vacuum region between the slabs are minimized. About 10 to 20 Å are usually sufficient to fulfill the requirement. For such a geometry, any three-dimensional electronic structure method able to treat open structures can be used and thus this a frequently employed model. Practical applications of these approaches are limited by the number of atoms in the three-dimensional supercell. Thus, a compromise needs to be found between slab thickness, space between the slabs, and the computational effort.
One way to overcome at least one of these limitations is the use of a single-slab geometry (cf. Fig. 3c)). The slab still has to be thick enough to achieve bulk-like behavior in its interior, but the correct vacuum boundary conditions of the semi-infinite vacuum on both sides of the slab are fulfilled. Besides the more accurate description of the vacuum, the surface state and the work-function, to computational effort may be reduced to 50% of the effort required in an supercell approach of a repeated slab model.
Finally, surface can be modeled by finite clusters (cf. Fig. 3d)). This approach has been widely used for the investigation of chemisorption, since it allows the application of standard quantum chemistry programs. While reasonable structural information such as adsorption geometries can be obtained with relatively small clusters consisting of 10 or 20 atoms, much larger clusters of preferably well over 100 atoms are required to achieve reliable results for sensitive quantities such as adsorption energies or the distinction between different adsorption sites with similar energy. However, even for large clusters, termination effects can have unpredictable side effects.

3 Pseudo-potentials and all-electron methods

In this section we will discuss the fundamental difference between the so called all-electron methods and the pseudo-potential methods. The all-electron methods do not modify the potential but have to cope with the $1/r$ singularity arising from the Coulomb attraction of the nuclei. The pseudo-potential approach on the other hand aims at generating a potential that replaces the singularity with a smooth potential while keeping the essential scattering properties. This then allows the use of relatively simple basis sets. We will come back to the discussion of the different basis sets in the next section.
In addition to the effect of the singularity onto the electrons relevant for the chemical binding, the presence of the singularity also induces states tightly bound to a single nucleus. These so called core-states I will discuss first before I introduce different pseudo-potential approaches.

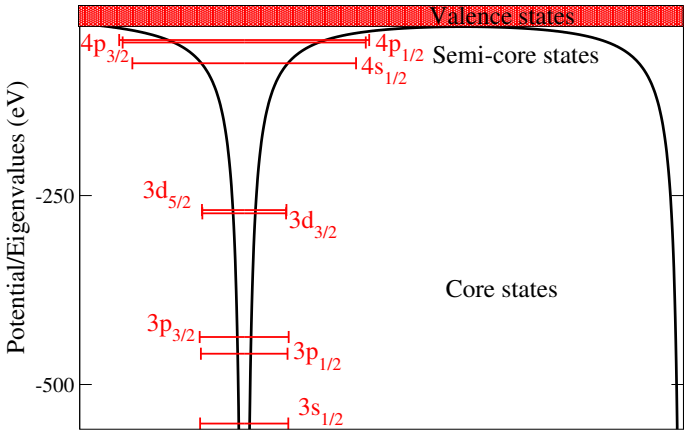


Fig. 4: *Schematics of the core potential and the eigenvalues. The bars indicate the (mean) extend of the wavefunction. The lowest 1s,2s,2p states are much deeper and not shown.*

3.1 Separation of core and valence states

While the eigenvalue problem Eq. (6) is solved for all occupied states i one can typically apply different boundary conditions to the low energy and high energy solutions. As shown schematically in Fig. 4 one distinguishes core electrons from valence electrons. The former have eigenenergies which are at least a couple of eV below the Fermi energy, the potential they experience is dominated by the coulomb attraction of the positively charged nucleus and to an excellent approximation spherically symmetric. The low energy solutions in this potential are confined to a space close to the nucleus and thus the wavefunctions have no overlap to neighboring atoms. The eigenvalue problem of these states can easily and up to high accuracy be solved applying the boundary conditions of isolated atoms, which is numerically tackled by a shooting method. Obviously, these core states do not participate in the formation of the chemical bond directly and they will only be weakly influenced by the potential variations introduced by the charge density due to the wavefunctions with higher energy not strongly localized at the atom cores.
The high energy, i.e. with energies close to the Fermi energy, valence electrons on the other hand are in general delocalized and can no longer be attributed to single atoms. It is the description of these electrons that require most computational effort. They form molecular orbitals in insulated systems or bands in crystalline solids. Some materials contain chemical elements with states (e.g. $5p$ states of $4f$ elements or W, p states of early transition metals) intermediate between valence and core states and those are coined semi-core states. These are high-lying and extended core states and particular care has to be taken on their treatment since their treatment as core states can cause significant errors in total energy, force and phonon calculations. According to the different treatment of the electrons, we decompose the charge density in the valence, semi-core and core densities

$$n(\mathbf{r}) = n_{\text{val}}(\mathbf{r}) + n_{\text{sc}}(\mathbf{r}) + n_{\text{core}}(\mathbf{r}), \tag{27}$$

the latter being spherically symmetric. The charge densities are calculated according to Eq. (1). Wavefunctions and energies of core states give access to hyperfine quantities such as isomer shifts, hyperfine fields and electric field gradient as well as chemical shifts of core levels.
This general discussion points to the strategy employed in density function theory calculations. The core states are calculated by solving an essentially atomic problem, frequently including relativistic effects by using a Dirac equation form of the Kohn-Sham equations. The valence

states are calculated separately by solving the Kohn-Sham equations in all space. The semi-core states have to be looked at carefully and either they are also included in the valance states calculation which make this calculation much more difficult as these states are relatively localized and very different from other valence states, or one includes the semi-core states into the core calculation, in which case one has to be careful not to confine the states artificially.

3.2 Pseudo-potentials

I will now give a short and superficial introduction into the large field of pseudo-potential methods. We start with the observation that unfortunately, even for eigenstates with high energy, i.e. for the valence states, the resulting wavefunction close to the nuclei is not smooth and changes rapidly on short length-scales. This can be understood by a simple argument: the smoothness of the wavefunction is of course related to its derivative, rapidly changing wavefunctions have a high derivative and hence a large momentum $p(\mathbf{r}) = \psi^*(\mathbf{r})\left(-i\hbar\nabla\right)\psi(\mathbf{r})$ and a high kinetic energy density $T(\mathbf{r}) = \frac{p^2(\mathbf{r})}{2m}$. Due to the locality of the DFT potential, the potential energy density $E_V(\mathbf{r}) = \psi^*(\mathbf{r})V(\mathbf{r})\psi(\mathbf{r})$ in the vicinity of the nuclei must be compensated by a correspondingly large $T(\mathbf{r})$ and thus a very rapidly varying wavefunction. Such a wavefunction is of course difficult to describe by a basis set that is not specifically adjusted to deal with the singular $1/r$ potential. An alternative idea is to modify the potential into a form easier to tackle. This strategy of replacing the potential is the core idea behind the so-called pseudo-potential approach.

At first it might be surprising that one can replace the singular $1/r$ potential by a pseudo-potential and still obtain useful results. However, one can quickly realize that one can replace the potential inside a sphere around the atoms by a different potential without changing the eigenvalue or the wavefunction in the space outside of this sphere. If we define such a smooth wavefunction $\tilde{\psi}(\mathbf{r})$ which equals the true wavefunction outside a sphere defined by a cutoff-radius, i.e. for $r > r_{cut}$, and expand it in the sphere into spherical harmonics and radial functions $\tilde{\psi}_l(r)$ we can use the equation

$$\left(-\frac{1}{2}\frac{d^2}{dr^2} + \tilde{V}(r) + \frac{l(l+1)}{2r^2}\right) r\tilde{\psi}_l(r) = \epsilon r\tilde{\psi}(r)$$

to define the potential $\tilde{V}(r)$, i.e. the potential is simply given by

$$\tilde{V}(r) = \left(r\tilde{\psi}_l(r)\right)^{-1}\left(\frac{1}{2}\frac{d^2}{dr^2}r\tilde{\psi}_l(r)\right) + \epsilon - \frac{l(l+1)}{2r^2}.$$

Obviously, this procedure requires that the smooth function $\tilde{\psi}_l$ was constructed without any nodes so that it can be inverted. Additionally, the function should behave as r^{l+1} for small r to avoid a singularity in $\tilde{V}$.

Replacing the true potential $V(\mathbf{r})$ by the potential defined by the $\tilde{V}_l$ leads to a Kohn-Sham equation with the same eigenvalue and the smooth eigenfunction $\tilde{\psi}(\mathbf{r})$. However, this very simple procedure already shows the various challenges in constructing such pseudo-potentials. First, the resulting potential is no longer a local potential. In fact, as we constructed a different smooth potential $\tilde{V}_l$ for each l-component, the total potential would have a form of

$$\tilde{V}(\mathbf{r}) = \sum_{l,m} \tilde{V}_l(r)|Y_{lm}\rangle\langle Y_{lm}|$$

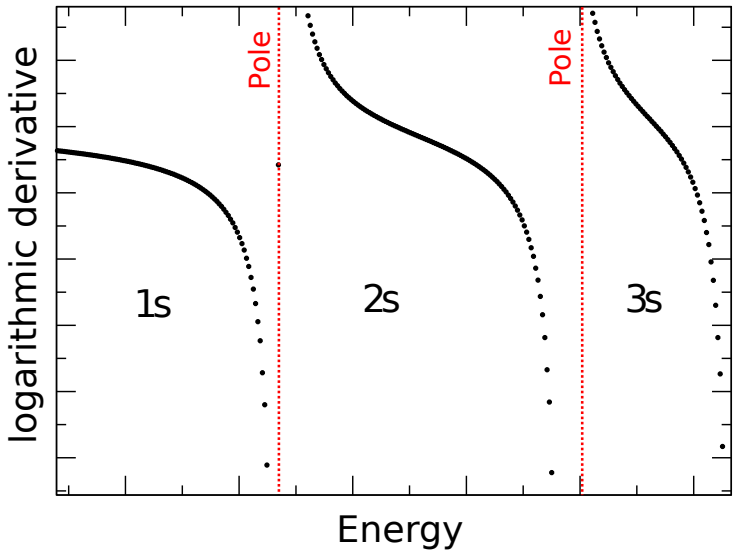


Fig. 5: *Logarithmic derivative, i.e.* $\left.\frac{\partial\psi_l(r)/\partial r}{\psi_l(r)}\right|_{r=r_0}$ *as a function of energy at some fixed radius* r_0*. Whenever the radial part of the wavefunction* $\psi_l(r)$ *has a node at* r_0 *a pole appears. These poles separate the different branches labeled* $1s$, $2s$, $3s$.

in which I used the Dirac notation for the non-local projector on the spherical harmonics $|Y_{lm}\rangle\langle Y_{lm}|$. While this of course adds additional complications to the solution of the Kohn-Sham equation, it is generally not avoidable as pseudo-potentials constructed in a purely local form usually only provide poor approximations. Even more significantly, one can conclude from our discussion that the potential $\tilde{V}(r)$ is constructed to provide the correct wavefunction for $r > r_{cut}$ only for a single eigenvalue. This is a serious restriction which can not be easily overcome. In fact, a pseudo-potential always can reproduce the correct eigenvalues exactly a very small (typically 1 or 2) discrete set of energies only and one has to ensure that even for other energies close by a good approximation is achieved. Finally, the construction of the pseudo-potential usually takes the spherically symmetric potential of an atom as a reference and the question how accurate the pseudo-potential approach can model the non-spherical potential contributions in a realistic material arrises.

On the other hand, our simple example also shows that we have a significant freedom in the construction of the pseudo-potential as the smooth wavefunction we started with was chosen with the only constraint of smoothness. Hence, a full zoo of different pseudo-potentials have been constructed ranging from empirical potentials which are constructed by fitting calculations using parametrized potentials to experimental results to potentials constructed using the projections of the span of the deep-lying core states to methods closer to our simplified discussion using the true atomic wavefunction.

Besides the fact that the pseudo-potentials are non-local and work only well in a limited range of energy, one also has to consider the norm of the wavefunctions. If one relaxes the requirement that the smooth wavefunction $\tilde{\psi}$ and the true wavefunction ψ have the same normalization one can usually create smoother pseudo-potentials. However, the charge difference in the sphere around the nuclei

$$\Delta q = e \int_{r<r_{cut}} |\psi(\mathbf{r})|^2 - |\tilde{\psi}(\mathbf{r})|^2$$

has to be accounted for. In the language of linear algebra this change in normalization leads

to the introduction of an overlap operator acting as a metric in the normalization and a generalized eigenvalue problem has to be considered. We will come to such problems later again when discussing the eigenvalue problem itself. Even though this leads to additional numerical complications, the advantages in smoothness usually make the use of non-local ultra-soft pseudo-potentials that are not norm-conserving attractive.
While very many different forms for pseudo-potentials have been suggested, constructed and successfully used to calculate a huge variety of systems, one must be aware of the fact that there is no single best pseudo-potential for all kind of calculations. In fact, the construction of a suitable pseudo-potential can be a painful process in which many reference calculations are needed to verify the good behavior of the potential. An important though very basic quantity to look at can be the logarithmic derivative of a single atom as shown in Fig. 5. While the pseudo-potential can not reproduce the same logarithmic derivative as given from the all-electron potential for all energies, it must be close to it in the energy range one wants to calculate eigenvalues.

3.3 PAW pseudo-potentials

A special form of pseudo-potentials was introduced by Blöchl[15] in the projector augmented wave (PAW) method. In contrast to standard pseudo-potential approaches this method allows to calculate in principle not only the smooth pseudo-wavefunctions but also the all-electron wavefunction and thus this method is not always considered a pseudo-potential approach. The basic idea of the method lies in the use of projector-function (hence the name) to deal with the non-smooth part of the wavefunction. It defines a transformation operator, which again is easily written using Dirac's notation as

$$\hat{O} = 1 + \sum_i \left(|\phi_i\rangle - |\tilde{\phi}_i\rangle \right) \langle p_i|$$

such that the projector function p_i localized in a sphere around the atom will be orthogonal to the smooth wavefunction $\tilde{\psi}_i$, i.e. $\langle p_i|\tilde{\phi}_j >= \delta_{ij}$. As a consequence, the transformation operator applied to the smooth wavefunction will replace the smooth part of the wavefunction in the spheres by the "true" wavefunction ψ_i. In practice, these "true" and smooth wavefunctions are constructed from a spherical atomic setup. Thus, one generally constructs a set of projectors p and corresponding wavefunctions ϕ and $\tilde{\phi}$ for the different atoms and the different $l-$quantum numbers relevant for the description of the valence states. The resulting PAW transformation operator than transforms a smooth pseudo-wavefunction into an all-electron wavefunction

$$\psi_{AE} = \hat{O}\tilde{\psi}.$$

The smooth wavefunction is obtained as a solution of a transformed Kohn-Sham equation in which the Hamiltonian and the overlap matrix enter which are obtained as

$$\hat{\tilde{H}} = \hat{O}^H \hat{H} \hat{O}$$

and

$$\hat{\tilde{S}} = \hat{O}^H \hat{O}.$$

The basis trick in this transformation lies in the use of a smooth projector function and in the replacement of the difficult true potential by a smooth pseudo-potential $\tilde{V}$. This is done by

transforming the potential operator as

$$\hat{\tilde{V}} = \hat{\tilde{V}} + \sum_{ij} |p_i\rangle \left(\langle \phi_i | \hat{V} | \phi_j \rangle - \langle \tilde{\phi}_i | \hat{\tilde{V}} | \tilde{\phi}_j \rangle \right) \langle p_j |.$$

These two steps ensure that the Hamiltonian $\hat{\tilde{H}}$ contains only terms that are easily expanded into smooth basis functions and that the same hold for the resulting smooth wavefunctions $\tilde{\psi}$. While the procedure used in defining the projection operator and the resulting transformations of the Hamiltonian are in principle exact mathematical transforms, in practice the resulting equations are only sufficiently smooth and reproduce the all-electron properties of the Hamiltonian with the true potential if both the projector functions and the smooth pseudo-potential $\tilde{V}$ are chosen carefully. PAW pseudo-potentials generally can be viewed as the most modern pseudo-potential approach applicable to the largest class of systems.

4 Basis sets

In this section two different approaches to solve the Kohn-Sham equation numerically will be discussed. First we spend some time on the discussion of different basis sets and later on I shortly introduce the Green function method which provides a different standpoint in solving the differential equation. While the introduction of a basis set might look like a rather simple and straightforward transformation of a differential equation into a linear algebra problem one should realize that it actually is very difficult to chose an appropriate basis set for a calculation. On the one hand the efficiency of the basis set is crucial in actual calculations, i.e. one wants to make sure that a small number of basis functions is sufficient to describe the wavefunctions accurately. The size of the basis set needed is of course one of the key parameters determining the feasibility of the calculation. In general, the memory required to solve the resulting eigenvalue problem scales quadratically in the basis size and the computational effort scales cubically. As the ultimately most efficient basis set in this respect would be the set of eigenfunctions one wants to calculate one strategy in choosing basis functions is to construct basis functions that approximate the final wavefunctions as best as possible without solving the eigenvalue problem already. For example one could use local solutions of the Kohn-Sham equation in some finite volume to expand the full Hamiltonian. We will discuss this idea shortly when discussing local basis sets and the LAPW method.

Another possibility in constructing basis sets is not to concentrate on the efficiency in the first place but to use a particularly simple basis set which enables quick calculations of the Hamiltonian and an efficient diagonalization by iterative procedures. The most prominent example of such a simple basis set is the plane-wave basis set discussed in the next section.

One should also note already here, that the choice of basis functions is also interlinked with further approximations in the solution of the Kohn-Sham equation. In particular, one should note that all-electron codes need other basis functions compared to pseudo-potential approaches and also the question whether one deals with periodic systems or insulated molecular setups make a difference when it comes to a suitable choice of basis functions.

4.1 Simple basis sets

Plane-waves

One of the most widely used basis sets are plane-waves

$$\phi_{\mathbf{g}}(\mathbf{r}) = \frac{1}{N} e^{i\mathbf{g}\mathbf{r}}$$

where the normalization N is related to the volume V of the system. Their popularity steams from a whole collection of useful properties which enable efficient calculations:

- The plane-waves form a set of orthonormal functions with

$$\int d^3r\, \phi^*_{\mathbf{g}}(\mathbf{r})\phi_{\mathbf{g}'}(\mathbf{r}) = \delta(\mathbf{g}' - \mathbf{g}).$$

- As they are eigenfunctions of the momentum operator, the kinetic energy is diagonal in these basis functions with eigenvalues $-\frac{1}{2}\mathbf{g}^2$.

- One can easily use the Bloch ansatz in periodic systems by choosing $\mathbf{g} = \mathbf{k} + \mathbf{G}$, with the $\mathbf{k}$-vector one wants to treat and $\mathbf{G}$ chosen as a reciprocal lattice vector. Then the expansion

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{g}} c_{\mathbf{g}} e^{i\mathbf{g}\mathbf{r}} = \frac{1}{N} e^{i\mathbf{k}\mathbf{r}} \sum_{\mathbf{G}} c_{\mathbf{k},\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}$$

  is a natural expansion of the Bloch state with a Fourier sum for the periodic function

$$u_{\mathbf{k}}(\mathbf{r}) = \frac{1}{N} \sum_{\mathbf{G}} c_{\mathbf{k},\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}.$$

- The potential matrix elements can be calculated as

$$\begin{aligned} \langle \mathbf{G}|\hat{V}|\mathbf{G}'\rangle &= \int \phi^*_{\mathbf{G}}(\mathbf{r}) V(r) \phi_{\mathbf{G}'}\, d^3r \\ &= \frac{1}{N^2} \int V(r)\, e^{i(\mathbf{G}'-\mathbf{G})\mathbf{r}} d^3r \\ &= V_{FT}(\mathbf{G}' - \mathbf{G}) \end{aligned}$$

  where $V_{FT}(\mathbf{G})$ is the Fourier transform of the potential $V(\mathbf{r})$. This Fourier transform can be calculated extremely efficiently using the FFT-algorithm. Furthermore, the matrix elements are independent of the $\mathbf{k}$-vector in question and thus would have to calculated only once.

Hence, using plane-waves, the Kohn-Sham equation transforms into a standard eigenvalue problem with the following Hamiltonian matrix

$$H_{\mathbf{G},\mathbf{G}'} = -\frac{1}{2}(\mathbf{k} + \mathbf{G})^2 \delta_{\mathbf{G},\mathbf{G}'} + V_{FT}(\mathbf{G} - \mathbf{G}').$$

This Hamiltonian matrix is usually a complex hermitian matrix, but in situations in which the system has inversion symmetry, i.e. $V(\mathbf{r}) = V(-\mathbf{r})$, the corresponding Fourier coefficients

$V_{FT}(\mathbf{G})$ are real and the linear eigenvalue problem simplifies to a real-symmetric problem. In light of these many advantages it cannot be surprising that many popular DFT codes are based on plane-wave basis sets (see e.g. Table 1). However, the plane-wave basis set has also some serious drawbacks. Most importantly, it is not easily possible to Fourier transform the $1/r$ potential. Its expansion decays only as $1/g^2$ and thus an extremely high number of $\mathbf{G}$ values would have to be used which makes plane-wave all-electron calculations impractical and requires the use of pseudo-potentials in plane-wave calculations. An additional drawback of plane-waves lies in the delocalization of the basis function. This results in a Hamiltonian matrix that is dense, i.e. a matrix in which all elements $H_{\mathbf{G},\mathbf{G}'} \neq 0$.

Local basis sets

Coming back to the idea of constructing a basis set that already tries to mimic properties of the final solution one arrives at the idea to expand the waverfunctions into localized functions that are solutions of an approximate local Kohn-Sham equation. Usually, one considers the spherically symmetric atomic problem for the construction of the basis and writes

$$\psi(\mathbf{r}) = \sum_{\mu} \sum_{i} c_{\mu,i} \phi_{\mu,i}(\mathbf{r} - \mathbf{r}_\mu)$$

where the different atoms are placed position $\mathbf{r}_\mu$ and at each of these atomic positions one uses a set of $\phi_{\mu,1}, \phi_{\mu,2}, \ldots$. The basis functions centered at different atoms have to overlap but this overlap can quickly vanish with distance, i.e. the functions $\phi_\mu(\mathbf{r})$ are centered around $\mathbf{r} = 0$, decay rapidly with increasing $\mathbf{r}$ and can be assumed to be zero at some point. Such a basis set again has many favorable properties:

- As the basis is explicitly atom centered, one can efficiently describe open systems with a large fraction of empty space such as molecules or surfaces with vacuum. Consequently, a lot of codes with a background in chemistry, focusing on molecular systems, are based on localized basis sets. One the other hand, periodic systems are not treated as naturally as in the case of plane-waves.

- The localization of the basis functions leads to sparse matrices. Since the Kohn-Sham Hamiltonian is local, matrix elements between basis functions belonging to atoms ν and μ located far apart vanish. For such basis functions one has

  $$0 = \int d^3r \, |\phi_{\nu,i}(\mathbf{r} - \mathbf{r}_\nu)||\phi_{\mu,j}(\mathbf{r} - \mathbf{r}_\mu)|$$

  and thus

  $$\begin{aligned} |H_{\nu,\mu}| &\leq \int d^3r \, |\phi_{\nu,i}(\mathbf{r} - \mathbf{r}_\nu)||\hat{H}(\mathbf{r})||\phi_{\mu,j}(\mathbf{r} - \mathbf{r}_\mu)| \\ &\leq |\max_{\mathbf{r}} \hat{H}(r)| \int d^3r \, |\phi_{\nu,i}(\mathbf{r} - \mathbf{r}_\nu)| \, |\phi_{\mu,j}(\mathbf{r} - \mathbf{r}_\mu)| = 0. \end{aligned}$$

- Beside functions like Gaussians, that are given analytically in terms of a few parameters, one also finds multiple incarnations of numerical orbitals. Following some optimization recipe, basis functions are constructed by numerical integration of a simplified or localized version of the Kohn-Sham equation. As these functions are derived from the Kohn-Sham equation itself, they already incorporate a lot of features the final wavefunctions

will have and thus provide an efficient basis set. Different strategies are available for such numerical orbitals, in some the basis set is adjusted to the potential in each convergence steps, in others the basis set is precomputed and kept fixed during charge self-consistency.

- As the basis functions are usually very efficient due to their centering at the atomic sites and the numerical optimization, one can achieve good accuracy with very small basis sets. As we shortly discussed in the introductory lecture (A1), the localized Wannier functions would provide the optimal localized basis set, in which one localized function per band, i.e. per valence electron, would be sufficient. To be as close as possible to this optimal basis set one has to find localized functions close to these Wannier functions.

One of the biggest drawbacks of localized basis sets, especially of numerical orbitals, is their unsystematic convergence behavior. It is very difficult to determine the level of accuracy the basis provides. Other basis sets, e.g. the plane-waves I discussed before, allow to increase the number of basis functions systematically to study the convergence. Since the local basis sets do not offer such a systematic sampling of the function space, this approach usually cannot be used to estimate the accuracy achieved and therefore a lot of experience is needed to ensure a proper choice of the basis functions. An additional complication in such codes arises from the fact that not only the wavefunction itself must be evaluated but also its square. As the products of basis functions occurring in this process are not necessarily included in the original set of localized functions, the charge density evaluated from the squares of the wavefunctions cannot be expressed in the same set of basis functions.

Real space grid

In some sense, the extreme point of a localized basis set is given if one attempts to expand the wavefunction on a discrete real-space grid[2]. While this idea of course is very attractive in terms of the ease of representation of quantities like the wavefunction, the charge density or the potential, it has a major drawback in the representation of the kinetic energy. The differential operator of the kinetic energy has to be approximated by a finite difference expression and thus the spacing of the grid and the order of this approximation plays an important role in these approaches. In principle the real-space approach and the plane-wave approach could be considered to have very similar properties as the Fourier transformation converts between these two representations. However, this would lead to an expression for the kinetic energy operator which if fully nonlocal on the real-space grid and thus not feasible in actual implementations. The analogy with the plane-wave basis set nevertheless holds in many aspects. The real-space grid approach also will only be able to work in a pseudo-potential framework and it shares the systematic convergence properties of the plane-wave methods, where one controls the accuracy of the basis by the grid spacing instead of the number of plane-waves. Main advantage of the real space grid approach lies in the extreme sparsity of the Hamiltonian and the favorable schemes for parallelization which are available (cf. Sec. 5.2).
Beside regular grids one could also consider grids which are finer in areas in which a better resolution of the wavefunction is needed. This idea leads to the finite element methods, which are extremely popular in engineering applications but so far not very widespread in the density functional theory community, and to methods which are based on wavelets.

[2]Depending on the exact form of the approach, the real-space grid representation can or cannot be viewed as an expansion into a basis set.

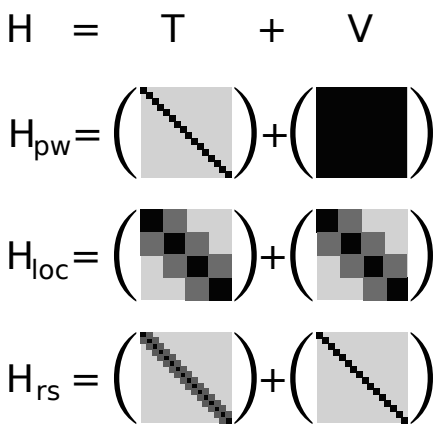


Fig. 6: *Typical form of matrices in a simple 1D setup. While a plane-wave (pw) basis set will diagonalize the kinetic energy, the real-space (rs) approach diagonalizes the potential term. The localized basis set (loc) will be sparse due to the finite range of the basis functions. The sparsity of the kinetic energy in the real-space approach is a consequence of a finite order finite difference approximation. Non-local operators as they appear in some pseudo-potentials are not considered here.*

Resulting eigenvalue problem

The different basis sets presented so far lead to different kind of eigenvalue problems. Most schemes produce a generalized eigenvalue problem either because of the non-orthogonality of the basis set itself, or by the use of a non-normconserving scheme in the pseudo-potential construction. Hence, two further basic properties of the eigenproblem are of key interest: the size of the matrix equations and the form of the matrices. While the size of the matrix of course depends on the system size, the required accuracy and the efficiency of the basis functions, the form of the matrix differs strongly depending on the basis set.
In general, the two major contributions, the kinetic energy and the potential term, behave differently. For example, in a plane-wave basis set the kinetic energy is given by a diagonal matrix and the potential matrix is a dense matrix with (in general) all elements of the matrix non-zero. The real-space basis provides a diagonal matrix for the potential. Fig. 6 shows the resulting form of the matrices for a simple one-dimensional example. The resulting form of the full Hamiltonian (and of the overlap matrix) can lead to substantial differences in the algorithms chosen for diagonalization, most obviously the question if a direct or an iterative solver is most appropriate and if special techniques for sparse matrices can be used. These numerical algorithms for matrix diagonalization are the theme of the manuscript D3 by E. Di Napoli.

4.2 APW-like concepts in all-electron calculations

The basis functions I discussed so far (with a few exceptions for some localized basis functions) can only be used together with the pseudo-potential approximation. We will now introduce a method for all-electron calculations. One typically works here with basis functions, which are the numerical solution of $(-\Delta + V_{\text{eff}} - E_l)\varphi = 0$ of the effective (spherical) potential containing the $1/r$ singularity, computed in a sphere around the atom. Therefore, these basis functions treat the singularity exactly. The matching of this wavefunction in such a sphere to the rest of the crystal outside the sphere divides the all-electron methods with regard to the eigenvalue dependence of the basis set into two groups: The nonlinear methods as for example

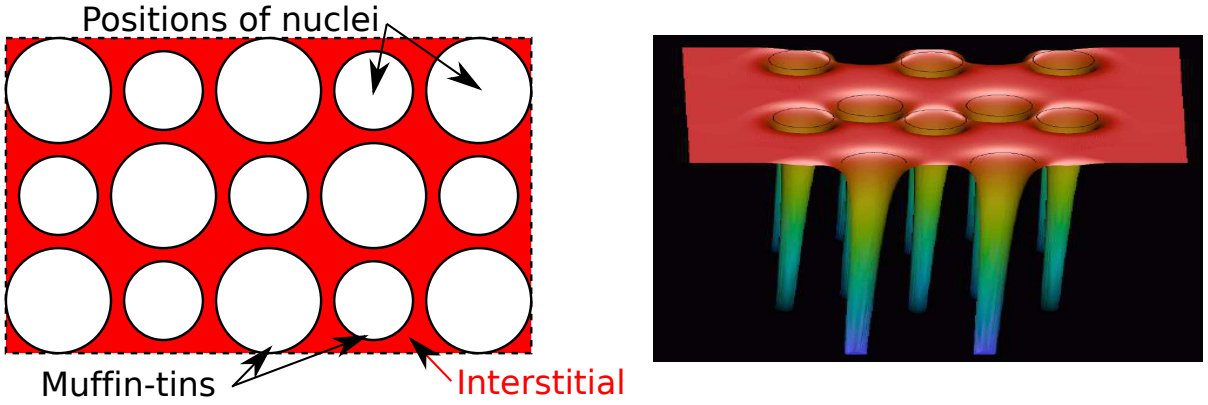


Fig. 7: *Left: Unit cell partitioned into (muffin-tin) spheres around atomic positions (two different atom kinds are assumed here with different radii) and the remaining interstitial volume. Right: Actual all-electron potential.*

the Korringa, Kohn and Rostocker (KKR) method and the APW method, and the linear methods, of which the most commonly used are the linear muffin-tin orbital method (LMTO) [16], the augmented spherical [17] and the APW-based schemes, e.g. FLAPW method.
In this section, I introduce step-by-step the full-potential linearized augmented plane-wave (FLAPW) method [18, 19], to solve the density-functional equations for a crystalline solid. The method originates from the APW method proposed by Slater [20, 21, 22]. Great progress of the APW methodology was achieved as the concept of linear methods [23, 16, 24, 25, 26], was introduced by Andersen and first applied by Koelling and Arbman using a model potential within the muffin-tin approximation. The linearized APW (LAPW) method reconciled the linear-algebra formulation of the variational problem with the convergence properties of the original formulation and allowed a straight forward extension of the method to the treatment of crystal potentials of general shape. The treatment of the potential and charge density without shape approximation [27, 28] and the implementation of the total energy [19] let to the development of FLAPW bulk [18, 28, 29, 30, 31, 32, 33, 34] film codes [18, 34, 35, 36]. It was during this time that the power and accuracy of the method were demonstrated to the community, largely through a series of calculations of surface and adsorbate electronic structures (for a review see Wimmer *et al.* [37]). These and other demonstrations established the FLAPW method as the method of choice for accurate electronic structure calculations for a broad spectrum of applications.
Constant conceptual and technical developments and refinements such as the proposal and implementation of the scalar-relativistic approximation (SRA) [38], the spin-orbit interaction by second variation [42], and the possibility to calculate forces [43, 44] acting on the ions to carry out structure optimizations, the proposal of a new efficient basis sets, the LAPW+LO [45] and APW+lo [46] basis, in which the APW basis is amended by local orbitals (lo) has made APW-like methods, and for our discussion the FLAPW method, a robust, versatile and flexible method, at reasonable computational expense.

The APW Concept

In the APW method the space is partitioned into spheres centered at each atom site, the so-called muffin-tins (MTs), and into the remaining interstitial region (cf. Fig. 7). The MT spheres do not

overlap and they are typically chosen such that they nearly (to allow for structural relaxations) fill the maximal possible space. Inside the muffin-tins, the potential is approximated to be spherically symmetric, and in many implementations the interstitial potential is set constant. The restrictions to the potential are commonly called shape-approximations. Noting that plane-waves solve the Schrödinger equation in a constant potential, Slater suggested to replace the Bessel functions $j_l(Kr)$ in the Rayleigh decomposition of the plane-wave inside the sphere by radial functions $u_l(K, r)$, which match the Bessel functions in value at the sphere radius R_{MT} and whose product with the spherical harmonics $Y_L(\hat{r})$ are the solutions in a spherical potential. It is this procedure what is understood by the term augmentation. Thus, the single wavefunctions $\psi_{k,\nu}(r)$ are expressed in terms of the APW basis functions:

$$\varphi_G(k,r) = \begin{cases} e^{i(\mathbf{k}+\mathbf{G})\mathbf{r}} & \text{interstitial region} \\ \sum_{lm} a_L^{\mu\mathbf{G}}(\mathbf{k}) u_l(r^\mu|E) Y_L(\hat{\mathbf{r}}^\mu) & \text{muffin-tin}\mu \end{cases} \tag{28}$$

The position $\mathbf{r}^\mu$ inside the spheres μ located at τ^μ (cf. Fig. 7) is given with respect to the center of each sphere. L abbreviates the quantum numbers l and m and u_l is the regular solution of the radial Schrödinger equation

$$\left\{ -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial r^2} + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2} + V(r) - E \right\} r u_l(r) = 0 \tag{29}$$

to the energy parameter E_l. Here, $V(r)$ is the spherical component of the potential $V(\mathbf{r})$. The coefficients

$$a_L^{\mu\mathbf{G}}(\mathbf{k}) = a_L^\mu(\mathbf{k}+\mathbf{G}) = 4\pi \exp(i\mathbf{k}\tau^\mu) i^l Y_L^*(\hat{\mathbf{K}}) \frac{j_l(KR^\mu)}{u_l(R^\mu)}, \qquad \mathbf{K} = \mathbf{k}+\mathbf{G} \tag{30}$$

are determined from the requirement, that the wavefunctions are continuous at the boundary of the muffin-tin spheres in order for the kinetic energy to be well-defined.
If E were kept fixed, used only as a parameter during the construction of the basis, the Hamiltonian could be set up in terms of this basis. This would lead to a standard secular equation for the band energies where for a given $\mathbf{k}$-point in the Brillouin zone (BZ) a set of band energies E_ν are determined. Unfortunately, it turns out, that the APW basis does not offer enough variational freedom if E is kept fixed. An accurate description can only be achieved if the energies are set to the corresponding band energies $\epsilon_{\mathbf{k},\nu}$. In this case the Hamiltonian matrix H depends not only on $\mathbf{k}$, $\mathrm{H}(\mathbf{k})$, but also on $\epsilon_{\mathbf{k},\nu}$, $\mathrm{H}(\epsilon_{\mathbf{k},\nu})$, and the latter can no longer be determined by a simple diagonalization. Since the u_l's depend then on the band energies, the solution of the secular equation becomes a nonlinear problem, which is computationally much more demanding than a secular problem. One way of solving this problem is to fix the energy E and scan over $\mathbf{k}$ to find a solution, i.e. find one band at the time, instead of diagonalizing a matrix to find all the bands at a given $\mathbf{k}$. Thus, in Slater's formulation of the method E enters as an additional non-linear variational parameter varying the shape of the functions u_l till the optimal shape is found for the band energies $\epsilon_{\mathbf{k},\nu}$ one has looked for. There are several other limitations connected to the APW method. One is rather obvious, when $u_l(R)$ in Eq. (30) becomes zero at the MT boundary, the radial function and the plane-wave becomes decoupled, known as the asymptote problem. Others are beyond the scope of the chapter. Further information about the APW method can be found in the book by Loucks [22], which also reprints several early papers including Slater's original publication [20].

There is one remaining point. Please notice that the APW method produces per construction principle wavefunctions with a discontinuity in the slope at the muffin-tin boundary. Due to these discontinuous first derivatives the secular equation in the APW basis

$$\sum_{\mathbf{G}'}(\langle\varphi_{\mathbf{G}}|\mathrm{H}-\varepsilon_{\mathbf{k}\nu}|\varphi_{\mathbf{G}'}\rangle+\langle\varphi_{\mathbf{G}}|\mathrm{T_S}|\varphi_{\mathbf{G}'}\rangle)\mathbf{c}^{\mathbf{G}'}_{\mathbf{k}\nu}=0 \tag{31}$$

contains a second term due to the matrix elements $\langle\psi|-\nabla^2|\psi\rangle$ of the kinetic energy operator T commonly defined as $\mathrm{T}=-\nabla^2$, which is replaced by $\langle\nabla\psi|\nabla\psi\rangle$, leading then via Green's theorem to the appearance of additional surface integrals $T_\mathrm{S}\propto\int\psi^*\left[\left(\frac{\partial\psi}{\partial n}\right)_--\left(\frac{\partial\psi}{\partial n}\right)_+\right]dS$, where $+(-)$ indicates just outside and inside the muffin-tin sphere. The matrix elements of $\mathrm{T_S}$ are proportional to the difference of the logarithmic derivatives from the function u_l, $D(u_l|E)=\frac{u_l'(R)}{u_l(R)}$, and that of an empty sphere $D(j_l|E)=\frac{j_l'(R)}{j_l(R)}$, taken at the sphere boundary. The logarithmic derivatives are related to the phase shifts in scattering events. Thus, the second term in Eq. (31) can be interpreted describing the scattering of a plane-wave coming from the crystal at the sphere of the atoms. It is well-known that the logarithmic derivatives and the phase shifts are energy dependent quantities, which explains the explicit energy dependence of the APW Hamiltonian in particular, and all nonlinear electronic structure methods in general.

The LAPW Basis-functions

To avoid the problems connected with the APW method resulting from the energy dependence of the Hamiltonian, in the middle of the seventies linearized methods were invented by Andersen [16] and Koelling and Arbman [24]. Based on an idea proposed by Marcus [26], the basis functions u_l in the muffin-tins were supplemented by their energy derivatives $\dot{u}_l$, but both, u_l and $\dot{u}_l$, are now evaluated at a fixed energy E_l. The original energy dependence of the radial basis-function is thereby replaced by the Taylor series:

$$u_l(E)=u_l(E_l)+(E-E_l)\dot{u}_l(E_l)+\cdots \tag{32}$$

terminated after the linear term. In this way, the wavefunctions are affected by an error which is quadratic in the deviation of the eigenvalue E from the energy parameter E_l, the error in the eigenvalues enter only to fourth order [24]. With this extension, the explicit form of the basis functions is now:

$$\psi_{\mathbf{G}}(\mathbf{k})=\begin{cases}\exp(i(\mathbf{k}+\mathbf{G})\mathbf{r}) & \text{interstitial}\\ \sum_{l,m}\left(a^{\mu,\mathbf{G}}_{lm}(\mathbf{k})u^\mu_l(r^\mu)+b^{\mu,\mathbf{G}}_{lm}(\mathbf{k})\dot{u}^\mu_l(r^\mu)\right)Y_{lm}(\hat{\mathbf{r}}^\mu) & \text{muffin-tin } \mu.\end{cases} \tag{33}$$

Examples of LAPW basis functions are shown in Fig. 8. The values of the coefficients $a^{\mu,\mathbf{G}}_{lm}(\mathbf{k})$ and $b^{\mu,\mathbf{G}}_{lm}(\mathbf{k})$ are determined to ensure continuity in value and derivative of the basis functions across the muffin-tin boundary. Thereby, also the surface integrals $\int\psi^*\left(\frac{\partial\psi}{\partial n}\right)dS$ which were encountered in the APW method disappear. In this way, the energy dependence of the Hamiltonian is removed, simplifying the eigenvalue problem Eq. (12), to a standard problem of linear algebra. Instead of working with u_l and $\dot{u}_l$ several LAPW implementations follow the APW idea, working only with u_l but for two different energy parameters E_l and E_l'. As we see below working with u_l and $\dot{u}_l$ is rather elegant.

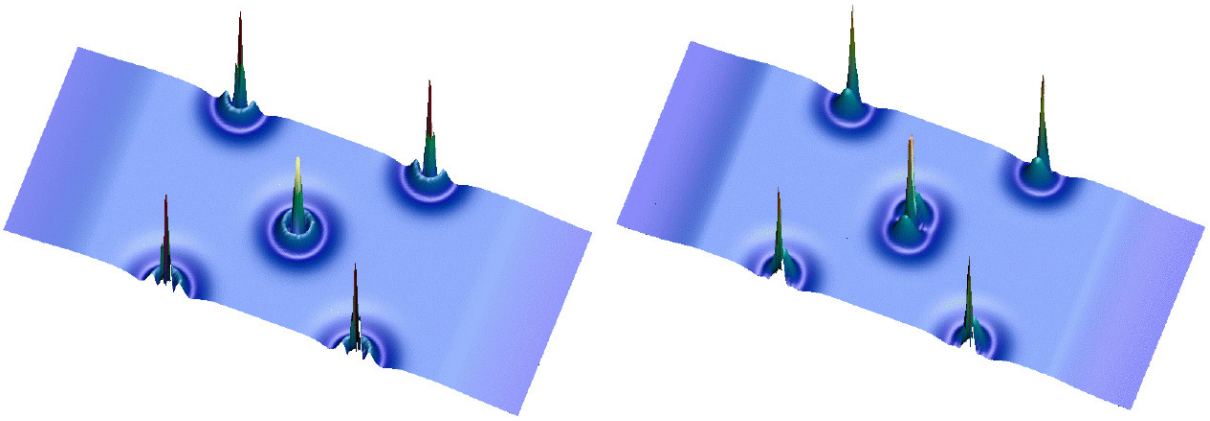


Fig. 8: *LAPW basis-function. Actually, the square* $|\phi_{\mathbf{k}}(\mathbf{r})|^2$ *for a* $\mathbf{k}$*-vector from the* $\overline{\Gamma}$*-point (left) and the* $\overline{M}$*-point of the two-dimensional Brillouin zone of a three layer film calculation. Note the strong variation of the wavefunction near the nucleii and the different character (*$4s$ *versus* $4p$*).*

If $\hat{H}^{\mu}_{sp}$ denotes the spherical Hamiltonian in Eq. (29), $\dot{u}$ can be determined from the energy derivative of this equation at E_l:

$$\hat{H}^{\mu}_{sp}\dot{u}^{\mu}_{l} = E_l\dot{u}^{\mu}_{l} + u^{\mu}_{l}. \tag{34}$$

The normalization of the radial functions is usually chosen like: [3]

$$\int_0^{R^{\mu}} r^2 {u^{\mu}_{l}}^2 dr = 1 \tag{35}$$

and the energy derivatives, $\dot{u}^{\mu}_{l}$, are orthogonal to the radial functions, i.e.

$$\int_0^{R^{\mu}} r^2 u^{\mu}_{l}\dot{u}^{\mu}_{l}\, dr = 0 \tag{36}$$

a relation, which will simplify the calculation of the elements of the Hamilton matrix.
Stimulated by the idea of the LAPW basis set, one may ask to improve the basis set by matching not only the first derivative continuously, but also higher derivatives working with higher energy derivatives of u_l. This approach has actually been followed by Takeda and Kübler [47] using n energy parameters to match the wavefunction continuously till the $(n-1)$st derivative. However, it turned out that such wavefunctions are variationally very stiff and the convergence of the results with respect of the number of basis functions is rather slow. This can be understood by following this procedure up to the extreme were the wavefunction matches to all derivatives. Then we know, the u_l must be the Bessel function j_l or the plane-wave, respectively. We have already argued before that this requires an infinite number of plane-waves to describe the wavefunction at the $1/r$ singularity. Thus, it is a great merit of the LAPW basis set, that the basis set is linear, but nearly as efficient as the APW method. The speed of convergence with

[3] In many LAPW-codes, the electrons in the muffin-tin are treated in the scalar-relativistic approximation [49]. This means that a two-component wavefunction is used and the normalization conditions are modified accordingly. For the continuity conditions, only the "large component" of the radial function is taken into account. To keep the formalism as simple as possible, in the following I will discuss only the non-relativistic case.

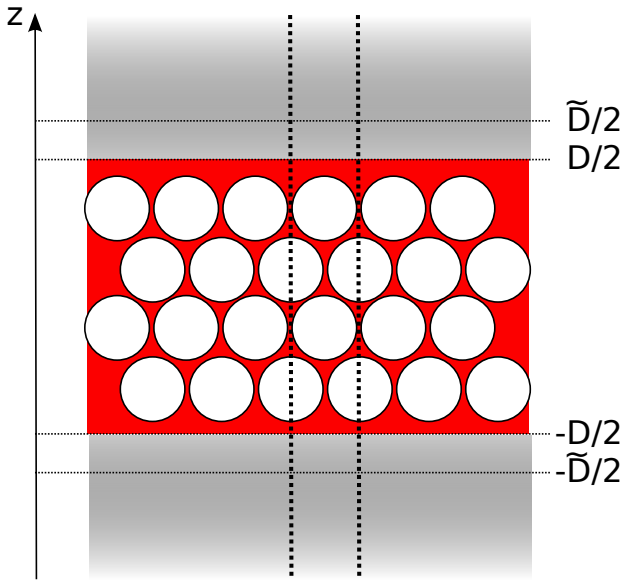


Fig. 9: *Setup used in a film calculation using the FLAPW method. In addition to the muffin-tin spheres and the interstitial two semi-infinite vacua are added.*

respect to the number of basis functions can even be improved by the introduction of local orbitals.

The energies E_l are chosen to minimize the linearization errors, i.e. in the center of gravity of the l-like bands. It should be noticed here, that the choice of the energy parameter in a certain sense also determines the nodal structure of the wavefunction. A basis function, where the $l = 1$ energy parameter is chosen to describe a $2s$-like wavefunction in a certain muffin-tin, will not be suitable to describe a $3s$ or a $1s$ state. The energy parameter is then said to be within the $2s$ branch (cf. Fig. 5). The flexibility of the basis function of course also depends on the size of the muffin-tin radius R, so that with the choice of a smaller R in some cases two branches can be forced to "collapse" to a single branch [48]. On the other hand, a smaller flexibility allows to separate core- from valence states in a calculation. Thus, in a typical calculation only high-lying valence states are calculated (e.g. $3s$, $3p$, $3d$), while very localized states (e.g. $1s$, $2s$, $2p$) are excluded from the calculation.

As a final point, I will address the question how large l should be in a realistic calculation. Since the a and b coefficients in Eq. (33) should ensure continuity across the muffin-tin boundary, the plane-wave cutoff, $G_{\max}$ and the l cutoff, $l_{\max}$, are normally chosen to match: A plane-wave with wave-vector $G_{\max}$ (given in inverse atomic units) has $G_{\max}/\pi$ nodes per atomic unit. A spherical harmonic with $l = l_{\max}$ has $2l_{\max}$ nodes along a great circle on the muffin-tin sphere, i.e. there are $l_{\max}/(\pi R)$ nodes per atomic unit. Therefore, a reasonable choice of the cutoffs is $l_{\max} = RG_{\max}$, typically $l_{\max} = 8$ is chosen.

The FLAPW-Method in Film Geometry for Surfaces and Thin Films

The idea of dividing space into different regions, i.e. into muffin-tins and the interstitial, can be generalized to treat surfaces very effectively in the single slab approximation (cf. Fig. 3c). One

now extends the space into three distinct regions, the muffin-tins, the interstitial and the vacuum region (cf. Fig. 9). The interstitial region now stretches from $-D/2$ to $D/2$ in z-direction, which is defined to be the direction perpendicular to the film. The representation of the wavefunctions inside the muffin-tin spheres remains exactly the same as in the bulk case. Since the periodicity along the z-direction is lost, the unit cell extends principally from $-\infty$ to ∞ in z-direction. Still the wavefunctions can be expanded in terms of plane-waves. However, the wave-vectors perpendicular to the film are not defined in terms of D, but in terms of $\tilde{D}$, which is chosen larger than D to gain greater variational freedom. Therefore, the plane-waves have the form

$$\varphi_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel, \mathbf{r}) = e^{i(\mathbf{G}_\parallel + \mathbf{k}_\parallel)\mathbf{r}_\parallel}\, e^{iG_\perp z} \qquad \text{with} \qquad G_\perp = \frac{2\pi n}{\tilde{D}}, \tag{37}$$

where $\mathbf{G}_\parallel$ and $\mathbf{k}_\parallel$ are the 2-dimensional wave- and Bloch vectors, $\mathbf{r}_\parallel$ is the parallel component of $\mathbf{r}$ and $G_\perp$ is the wave-vector perpendicular to the film. The basis functions in the vacuum region are constructed in the same spirit as the functions in the muffin-tins. They consist of plane-waves parallel to the film, and a z-dependent function $u_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z)$, which solves the corresponding one-dimensional Schrödinger equation Eq. (38), plus its energy derivative $\dot{u}_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z)$.

$$\left\{ -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + V_0(z) - E_{vac} + \frac{\hbar^2}{2m}(\mathbf{G}_\parallel + \mathbf{k}_\parallel)^2 \right\} u_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z) = 0 \tag{38}$$

E_{vac} is the vacuum energy parameter and $V_0(z)$ is the planar averaged part of the vacuum potential. As in the case of $\dot{u}_l$ in the muffin-tins, the function $\dot{u}_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z)$ is calculated from a Schrödinger-like equation, which can be obtained by deriving Eq. (38) with respect to the energy.

$$\left\{ -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + V_0(z) - E_{vac} + \frac{\hbar^2}{2m}(\mathbf{G}_\parallel + \mathbf{k}_\parallel)^2 \right\} \dot{u}_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z) = u_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z) \tag{39}$$

The resulting basis functions have the form

$$\varphi_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel, \mathbf{r}) = \left\{ a_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel) u_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z) + b_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel) \dot{u}_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel, z) \right\} e^{i(\mathbf{G}_\parallel + \mathbf{k}_\parallel)\mathbf{r}_\parallel} \tag{40}$$

The coefficients $a_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel)$ and $b_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel)$ are determined in exactly the same way as it is done for the muffin-tins by requiring that the functions are continuous and differentiable at the vacuum boundary. It should be mentioned, that the vacuum basis functions offer less variational freedom than the basis set in the interstitial region does. This can be seen by noting that there are only two functions, $u_{\mathbf{G}_\parallel}$ and $\dot{u}_{\mathbf{G}_\parallel}$ times the corresponding planar plane-wave, to be matched to all plane-waves of the interstitial region with the same $\mathbf{G}_\parallel$. But there are generally far more than two different $G_\perp$'s, i.e the number of basis functions in the vacuum region is significantly smaller than in the interstitial region. However, this can be improved rather easily. In Eq. (38) only one energy parameter E_{vac} is used. Instead one can use a whole series of parameters E^i_{vac} to cover an energy region. A possible choice of the energy parameters could be $E^i_{vac} = E^{G_\perp}_{vac} = E_{vac} - \frac{\hbar^2}{2m}G_\perp^2$, which correspondingly leads to $G_\perp$ dependent basis functions $u_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel, z)$. For more details see Ref. [50]. In general, however, the present approximations is accurate enough due to the small weight of the wavefunctions in the vacuum region.

Finally I would like to summarize the basis set used for thin film calculation with the FLAPW

method.

$$\varphi_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel,\mathbf{r}) = \begin{cases} e^{i(\mathbf{G}_\parallel+\mathbf{k}_\parallel)\mathbf{r}_\parallel}\ e^{iG_\perp z} & \text{interstitial} \\ \Big\{a_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel)u_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel,z) & \\ \quad +b_{\mathbf{G}_\parallel G_\perp}(\mathbf{k}_\parallel)\dot{u}_{\mathbf{G}_\parallel}(\mathbf{k}_\parallel,z)\Big\}\, e^{i(\mathbf{G}_\parallel+\mathbf{k}_\parallel)\mathbf{r}_\parallel} & \text{vacuum} \\ \sum_L a_L^{\mu\mathbf{G}}(\mathbf{k})u_l(r)Y_L(\hat{\mathbf{r}}) + b_L^{\mu\mathbf{G}}(\mathbf{k})\dot{u}_l(r)Y_L(\hat{\mathbf{r}}) & \text{MT}^\mu \end{cases} \tag{41}$$

This expansion has been suggested by H. Krakauer, M. Posternak and A. J. Freeman [35].
The expansion of the charge density n and the potential is very similar to the expansion of the wavefunction. In the interstitial-region the two quantities are expanded into three-dimensional plane-waves, inside the muffin-tins they are represented by spherical harmonics and radial functions, which are stored on an exponential mesh and in the vacuum they are expanded into two-dimensional plane-wave and z-depended functions. Of course, the charge density and the potential exhibit the lattice symmetry. Therefore, the expansion into plane-waves is more general than necessary. The plane-waves can be replaced by symmetrized plane-waves, the so called stars Φ_s^{3D} for the interstitial region and the two-dimensional stars $\Phi_s^{2D}(\mathbf{r})$ for the vacuum region. Thus, the charge density and potential is expanded in the form:

$$n(\mathbf{r}) = \begin{cases} \sum_s n_s \Phi_s^{3D}(\mathbf{r}) & \mathbf{r} \in \text{interstitial region} \\ \sum_s n_s(z) \Phi_s^{2D}(\mathbf{r}) & \mathbf{r} \in \text{vacuum} \\ \sum_\nu n_\nu^\mu(r) K_\nu(\hat{\mathbf{r}}) & \mathbf{r} \in \text{MT}^\mu \end{cases} \tag{42}$$

and the Hamiltonian and overlap matrix consists now of three terms: the interstitial, muffin-tin sphere and the vacuum contribution, paying tribute that the space is now partitioned in three regions

$$\hat{H} = \hat{\mathrm{H}}_I + \hat{\mathrm{H}}_{MT} + \hat{\mathrm{H}}_V \qquad \text{and} \qquad \hat{\mathrm{S}} = \hat{\mathrm{S}}_I + \hat{\mathrm{S}}_{MT} + \hat{\mathrm{S}}_V. \tag{43}$$

4.3 The Green function methods

So far I discussed methods to obtain the solution of the Kohn-Sham equation in terms of the eigenfunctions of the Kohn-Sham Hamiltonian. An alternative way to solve the Kohn-Sham equation is provided by the calculation of the single particle Green function. While this approach has a couple of disadvantages like the need to do an energy integration and a more sophisticated determination of the Fermi-level, it also has distinct advantages, most of all the possibility to easily treat energy dependent quantities or non-linear approaches for all-electron methods. Hence, I will shortly discuss the general properties of the Green functions, the basics of the method due to Korringa, Kohn and Rostocker (KKR) to calculate the Green function and the embedding scheme which allows to obtain the Green function of a subsystem embedded into an outer volume.

The Single Particle Green Function

In density functional calculations the solution of the Kohn-Sham equations for the single particle wave functions $\psi_\alpha(\mathbf{r})$ and the corresponding eigenvalues ε_α, the single particle energies,

represents the central problem. Thus, most of electronic structure calculations follow this route, i.e. calculating eigenfunctions ψ_α and eigenvalues ε_α. However, the calculation of ψ_α and ε_α can be avoided, if instead the single particle Green function $G(\mathbf{r}, \mathbf{r}'; E)$ of the Kohn-Sham equation is determined, since this quantity contains all the information about the ground state. In particular the charge density and the local density of states can be directly calculated from the Green function, which is the solution of the Schrödinger equation for an energy E with a source at position $\mathbf{r}'$:

$$(-\frac{1}{2}\nabla^2 + V(\mathbf{r}) - E)\, G(\mathbf{r}, \mathbf{r}'; E) = -\delta(\mathbf{r} - \mathbf{r}') \quad . \tag{44}$$

Using the spectral representation for the (retarded) Green function

$$G(\mathbf{r}, \mathbf{r}'; E + i\epsilon) = \sum_\alpha \frac{\psi_\alpha(\mathbf{r})\psi_\alpha^*(\mathbf{r}')}{E + i\epsilon - E_\alpha} \tag{45}$$

it is easy to show that the charge density $n(\mathbf{r})$ can be directly expressed by an energy integral over the imaginary part of the Green function:

$$n(\mathbf{r}) = 1 \sum_{\substack{\alpha \\ E_\alpha < E_F}} |\psi_\alpha(\mathbf{r})|^2 = -\frac{1}{\pi} \int^{\epsilon_F} dE\, \mathrm{Im}\, G(\mathbf{r}, \mathbf{r}; E) \tag{46}$$

This relation directly allows the calculation of the charge density from the imaginary part of the Green function, which can be interpreted as the local density of states at the position $\mathbf{r}$.
In this way the evaluation of the wave-functions $\psi_\alpha(r)$ can be avoided. Due to the strong energy-dependent structure of the density of states, the evaluation of the energy integral is usually very cumbersome and typically about 10^3 energy points are needed in an accurate evaluation of this integral.
The numerical effort can be strongly decreased, if the analytical properties of the Green function $G(z)$ for complex energies $z = E + i\Gamma$ are used. Since $G(z)$ is analytical in the upper half of the complex energy plane, the energy integral can be transformed into a contour integral in the complex energy plane

$$n(\mathbf{r}) \;=\; -\frac{1}{\pi}\, \mathrm{Im} \int_{E_B}^{E_F} dz\, G(\mathbf{r}, \mathbf{r}; z) \tag{47}$$

where the contour starts at an energy E_B below the bottom of the valence bands, goes into the upper half of the complex plane and comes back to the real axis at the Fermi level. Since for complex energies all structures of the Green function are broadened by the imaginary part Γ, the contour integral can be accurately evaluated using rather few energy points, typically 20-30, leading to a large saving of computer time. In this way Green function methods are competitive with diagonalization methods. Additional advantages occur for systems with two- or three-dimensional symmetry, since as a result of the energy broadening the k–integration over the Brillouin zone requires for complex energies much less $\mathbf{k}$–points. In the evaluation of the contour integral, special care is necessary for the piece of the path close to ϵ_F, since here the full structure of $G(E)$ on the real axis reappears. Therefore, the energy mesh should become increasingly denser when approaching ϵ_F.
The integration over a complex energy contour can also be extended to finite temperatures by using the analytical properties of the Fermi-Dirac distribution. Here the essential point is that the contour close to ϵ_F is replaced by a sum over Matsubara energies $z_j = \epsilon_F + i\pi(2j-1)kT$,

$j = 1, 2, \ldots$. Then, only complex energies are needed, since the energy point closest to ϵ_F has still an imaginary part of πkT. This is of particular advantage, when a discrete k-mesh is used, like e.g. in the special points method.
The real problem is the evaluation of the Green function for the system of interest. Since we want to avoid evaluation of all eigenvalues ε_α and wave functions ψ_α we rather calculate the Green function G

$$G(E) = \frac{1}{E + i\varepsilon - H} = \frac{1}{E + i\varepsilon - H_o - V} \tag{48}$$

of a system with Hamiltonian $H = H_o + V$. We start by constructing the Green function $G_o = \{E + i\varepsilon - H_o\}^{-1}$ of a reference system, which is analytically known or easy to calculate. Then $G(E)$ can be obtained from the Dyson equation

$$G(E) = G_o(E) + G_o(E)\, V\, G(E) = G_o\, \frac{1}{1 - VG_o}\,. \tag{49}$$

For instance, for a bulk crystal one starts with the free space Green function $G_o(H_o = -\frac{1}{2}\nabla^2)$, such that V is the sum of the potentials of all atoms. For the surface Green function, G_o is identified with the bulk Green function, such that V is the difference between the potentials at the surface and in the bulk. Analogously for a cluster of adatoms on a surface one starts again with the surface Green function G_o, such that V represents the change of the ad-cluster potential with respect to the surface potential including the perturbation of the potentials of the neighboring host atoms. Most importantly the perturbed potential V is well localized near the impurities, while the perturbed wavefunctions are not localized and accurately described by the Dyson equation.
Such impurity problems are often described by an 'Ersatzgeometry', e.g. an impurity in a relatively small cluster of atoms or by a supercell geometry with a periodic array of impurities. In these cases the boundary conditions for the wave functions are changed violently, since e.g. for a cluster all wave functions are restricted to the size of the cluster. Therefore, the introduction of the host Green function G_o solves the so-called "embedding problem", since it correctly describes the embedding of the impurity in the infinite surface system.

KKR representation of the Green function

The multiple-scattering method of Korringa, Kohn and Rostoker (KKR) for the calculation of the electronic structure of materials was introduced in 1947 by Korringa [51] and in 1954 by Kohn and Rostoker [52]. A further significant development of the KKR scheme came when it was reformulated as a KKR Green function method [53, 54]. It has special advantages in many cases because the method is able to produce the crystal Green function efficiently by relating it to the Green function of free space via the Dyson equation. In a second step the crystal Green function can be used as a reference in order to calculate the Green function of an impurity in the crystal [55], again via a Dyson equation. This way of solving the impurity problem is extremely efficient, avoiding the construction of huge supercells which are needed in wavefunction methods. The development of *screened*, or *tight-binding*, KKR was a further breakthrough for the numerical efficiency of the method [56]. For some more information on these recent developments I refer to the contribution D6 of R. Zeller of this spring school.
In the KKR method the Schrödinger equation is solved by multiple scattering theory, describing the propagation of a wave in the solid as a repetition of single scattering events at the different atoms. Thus, first the single scattering event of the wave at the potential of the different single

atoms n is calculated, described by the single site "t-matrix" $t_{n'}$, and then the multiple scattering at the given arrangement of the atoms in the crystal. The resulting equations show a beautiful separation between potential and structural properties, which are typical for the KKR method. In the following I summarize the most important results.

In the KKR-Green function method one divides the whole space into non–overlapping and space–filling cells centered at positions r^n (similar to Fig. 7). In each cell the electrons are scattered by potentials v^n, which in this section are assumed to be spherically symmetric and centered at r^n. By introducing cell-centered coordinates the Green function $G(r+r^n, r'+r^{n'}; E)$ can then be expanded in each cell as a function of r and r' into spherical harmonics:

$$\begin{aligned} G(\mathbf{r}_n + \mathbf{r}, \mathbf{r}_{n'} + \mathbf{r}'; E) &= -i\sqrt{E} \sum_L R_L^n(\mathbf{r}_<; E) H_L^n(\mathbf{r}_>; E) \delta_{nn'} \\ &+ \sum_{LL'} R_L^n(\mathbf{r}; E) G_{LL'}^{nn'}(E) R_{L'}^{n'}(\mathbf{r}'; E) \end{aligned} \tag{50}$$

Here $\mathbf{r}$ and $\mathbf{r}'$ are restricted to the cells n and n' and $\mathbf{r}_<$ and $\mathbf{r}_>$ denote the one of the two vectors $\mathbf{r}$ and $\mathbf{r}'$ which has the smaller or larger absolute value. The $R_L^n(\mathbf{r}; E)$ and $H_L^n(\mathbf{r}; E)$ are the product of spherical harmonics and radial eigenfunctions to the central potential $v^n(r)$:

$$R_L^n(\mathbf{r}; E) = R_l^n(r; E)\, Y_L(\hat{\mathbf{r}}), \tag{51}$$

$$H_L^n(\mathbf{r}; E) = H_l^n(r; E)\, Y_L(\hat{\mathbf{r}}). \tag{52}$$

Here $R_L^n(r, E)$ is the regular solution which varies at the origin as r^l and which represents the solution for an incoming spherical Bessel function $j_l(\sqrt{E}r)Y_L(\hat{\mathbf{r}})$, while H_l^n is the corresponding irregular solution varying as $1/r^{l+1}$ at the origin and being identical with the spherical Hankel function $h_l(\sqrt{E}r)$ outside the range of the potential. Both radial functions are connected by the Wronskian relation, which guarantees that the first term in Eq. (50) represents the exact Green function for the single potential $v^n(r)$ in free space. Since this term satisfies already the source condition $-\delta(\mathbf{r} - \mathbf{r}')$ for the Green function of Eq. (44), the second term is source free and contains in the double angular momentum expansion only the regular solutions R_L^n and $R_{L'}^{n'}$.

By construction, the expression (50) for the Green function satisfies in each cell n the general solution of the Schrödinger equation (44) for the Green function, while the matrix $G_{LL'}^{nn'}(E)$, the so-called *structural Green function*, describes the connection of the solutions in the different cells and thus contains all the information about the multiple scattering problem, which is in this way reduced to the solution of an algebraic problem. The clear separation between the single–site properties, described by the radial solutions $R_L^n(r)$ and $H_L^n(r)$ and the multiple scattering properties as described by the matrix $G_{LL'}^{nn'}$, is the main advantage of the KKR method.

In principle, the structural Green function matrix $G_{LL'}^{nn'}(E)$ can be determined by matching the solutions of the neighboring cells at the cell boundaries. However at the cell boundaries the angular momentum expansion converges rather slowly, so that presumably a large $l_{\max}$ cut-off would be needed. The more elegant and at the same time more efficient way consists in using the power of multiple scattering theory, where the Green function is basically only needed in the inner region of the cell, where the potential is strong, so that the l–convergence does not represent any problem. As shown by Beeby and others [53], the structural Green function matrix can be determined from the corresponding matrix g in free space by the Dyson equation

$$G_{LL'}^{nn'}(E) = g_{LL'}^{nn'}(E) + \sum_{n''L''} g_{LL''}^{nn''}(E)\, t_{l''}^{n''}(E)\, G_{L''L'}^{n''n'}(E) \tag{53}$$

where the t-matrix t_l^n for the potential $v^n(r)$ is given by

$$t_l^n(E) = \int_0^R r^2\, dr\; j_l(\sqrt{E}r)\; v^n(r)\; R_l^n(r;E) \tag{54}$$

The derivation of this equation is lengthy and straightforward, so that I refer for this to the literature cited above. An elementary derivation, valid also for the full–potential case, has been given by Zeller [57].
In practice, the host structural Green functions are first calculated in k-space using matrix inversion; a subsequent Fourier transform gives us the real-space quantities. We write, then,

$$G_{LL'}(k;E) = \sum_{n'} G_{LL'}^{nn'}(E)\, e^{-i\mathbf{k}\cdot(\mathbf{r}^n - \mathbf{r}^{n'})} \tag{55}$$

(which, due to translational symmetry, is independent of n). The algebraic Dyson equation Eq. (53) becomes

$$G_{LL'}(k;E) = g_{LL'}(k;E) + \sum_{L''} g_{LL''}(k;E)\, t_{l''}(E)\, G_{L''L'}(k;E) \tag{56}$$

(the t-matrix is independent of n, again due to translational symmetry). Here $g_{LL'}$ are the reference structural green function of the original system before perturbing it by the surface. This original system can be for example free space. The structural Green functions $G_{LL'}$ and $g_{LL'}$, and the t-matrix t_l, are considered as matrices in L and L', and (56) is solved by matrix inversion after a cutoff at some $l = l$ for which the t-matrix becomes negligible (usually $l = 3$ or 4 suffices). The result is

$$G(E) = \frac{1}{V}\int d^3k\, e^{i\mathbf{k}\cdot(\mathbf{r}^n - \mathbf{r}^{n'})}\left[(1 - g(\mathbf{k};E)\, t(E))^{-1}\, g(\mathbf{k};E)\right]_{LL'} \tag{57}$$

where the integral is over the Brillouin zone volume V. For the calculation of the charge density or of the density of states, only the on-site term $n = n'$, $G_{LL'}^{nn}(E)$, is needed.
Here, the t-matrix $t(E)$ depends on the atom-type μ and on angular-momentum indexes (it is site-diagonal, $(t)_l^{\mu\mu'} = t_l^n\, \delta_{\mu\mu'}$). The structure constants $g(\mathbf{k};E)$ are considered as a matrices in both (L, L') and (μ, μ'), and thus the computational effort for the matrix inversion increases as $O(N^3)$. A considerable speed-up can be achieved for large systems by using the concept of the screening transformation.

The Green function Embedding method

As the Green function itself is an energy dependent quantity, it is the most natural quantity to calculate if one has to include energy dependent terms into the Hamiltonian. Such additional terms are usually called self-energy operators and in most general form are written as $\Sigma(\mathbf{r}, \mathbf{r}', \epsilon)$, i.e. these are non-local energy dependent operators. This self-energy is one of the most basic quantities in the quasi-particle theory and is for example calculated in the so-called GW-approximation (cf. manuscript A4 by Ch. Friedrich). Another situation in which a self-energy appears is the case of open systems, i.e. of systems which are only part of a larger "outside" and in which electrons can enter or leave from this outer volume. Typical examples of such geometries are transport setups in which the electrons enter from a lead, are scattered in

the scattering volume and leave via the same or a different lead again. Here the Green function of the scattering volume can be obtained by including a self-energy operator which describes the effect of the leads (cf. manuscript A9 by S. Tsukamoto).
Another name of such a self-energy containing the effect of an outside volume was coined by J. Inglesfield[58]: embedding potential. The question of how to obtain the embedding potential in a numerical scheme will not be discussed here, instead I refer to the corresponding literature[12]. Instead the focus should return to the Kohn-Sham equation with the additional embedding potential which reads in the approach of Inglesfield as

$$(\hat{H}(\mathbf{r}) - \epsilon)\psi(\mathbf{r}) + \frac{1}{2}\delta(n - n_s)\left[\partial_n\psi(\mathbf{r}) - 2\int_S \Sigma(\mathbf{r}, \mathbf{r}_s{}'; \epsilon)\psi(\mathbf{r}_S{}')\, dS'\right] = 0, \tag{58}$$

here $\hat{H}$ denotes the standard Kohn-Sham Hamiltonian and the addition term acts on the surface S of the volume considered only (indicated by the Delta function in which only the normal coordinate enters). The surface derivative of the wavefunction ∂_n is assumed to be the derivative along the surface normal. Eq. (58) can be written symbolically as

$$(\hat{H}_{\text{emb}}(\epsilon) - \epsilon)\psi(\vec{r}) = 0, \tag{59}$$

where the Hamiltonian $\hat{H}_{\text{emb}}$ has been defined to include the normal Hamilton operator as well as the surface derivative and the non-local embedding potential acting on the boundary surface. The embedding potential and therefore the embedding Hamiltonian $\hat{H}_{\text{emb}}$ is energy dependent and the direct solution of Eq. (58) for the wavefunctions is in general not possible[4]. Instead the most useful quantity to look at is the Green function. Using the modified embedding Hamiltonian H_{emb} the Green function is given by

$$(H_{\text{emb}}(\epsilon) - \epsilon)G(\vec{r}, \vec{r}\,'; \epsilon) = -\delta(\vec{r} - \vec{r}\,'). \tag{60}$$

From this Green function the charge in the embedded volume can be calculated or transport properties can be obtained. The embedding method also allows to efficiently treat surface in a semi-infinite surface model (cf. Fig. 3a)

5 A Zoo of Electronic Structure Methods

The discussion of pseudo-potentials, basis sets and Green functions can only give a brief glimpse on the enormous variety of available methods for DFT calculations. Basically, looking at the Kohn-Sham equation many different approximations and numerical approaches can be chosen to treat the different terms, Fig. 10 gives a graphical representation of this zoo of methods. The kinetic energy can be approximated by a non-relativistic treatment, the so called scalar relativistic approximation, one can additionally include spin-orbit interaction or even employ the full relativistic treatment by the Dirac equation (cf. the manuscript A10 by G. Bihlmayer). The potential term can also be approximated in multiple ways: the shape of the potential might be approximated, a pseudo-potential might be used or the full-potential all-electron approach can be chosen. The spin degree of freedom offers an additional range of different approximations (cf. manuscript C4 of S. Blügel). The different approximations for the exchange-correlation

[4] Actually, due to the need to specify the boundary condition of the wavefunction different embedding potentials exist and no easy method exists to construct the correct embedding potential for the wavefunction.

Code	Basis Set	Potential	Homepage
ABINIT	Plane-wave	Pseudo	www.abinit.org
BigDFT	Wavelet	Pseudo	bigdft.org
CASTEP	Plane-wave	Pseudo	www.castep.org
CP2K	Localized+Plane-wave	Pseudo	www.cp2k.org
CPMD	Plane-wave	Pseudo	www.cpmd.org
ELK	FLAPW	All-electron	elk.sourceforge.net
FLEUR	FLAPW	All-electron	www.judft.de
FHI aims	Localized	Pseudo	aimsclub.fhi-berlin.mpg.de
GPAW	Real space grid	Pseudo	wiki.fysik.dtu.dk/gpaw/
ONETEP	Plane-wave	Pseudo	www.onetep.org
Quantum ESPRESSO	Plane-wave	Pseudo	www.quantum-espresso.org
SIESTA	Localized	Pseudo	departments.icmab.es/leem/siesta/
VASP	Plane-wave	Pseudo	cms.mpi.univie.ac.at/vasp/
WIEN2K	FLAPW	All-electron	www.wien2k.at

Table 1: *Some DFT codes. Even though the list contains some of the most popular codes, it is far from complete and has a focus on european developments from groups participating in the Psi-k network (see http://www.psi-k.org).*

potential also increase the number of choices one can make in performing DFT calculations. Finally, also the more numerical aspects of which basis set to choose and which algorithm to use for the solution of the resulting differential equation sets the different DFT codes apart. Thus, it can be no surprise that many different codes are available (for a short list see Table 1) and one usually has to choose an appropriate method and code very carefully taking into account the approximations made and the accuracy and precision needed.

5.1 The CPU Time Requirement

The number of basis functions N is determined by the required precision P of a calculation and by the volume Ω of the unit cell or the number of atoms in the unit cell, N_A, respectively. For three-dimensional unit cells N usually scales as $N \propto P^3$. Also other parameters like the number of $\mathbf{k}$-points needed, depend on the required precision and the (inversely) on the system size. Hence, in general, the triple (N_k, M, N), the number of $\mathbf{k}$-vectors in the BZ used, the number M of states considered, and the number of basis functions N are determined by the required precision of the calculation and by the volume of the unit cell. These parameters determine the CPU-time and memory requirements of the calculations. Keeping the loop-structure in mind exhibited in Fig. 1, typically the calculation CPU time scales as

$$\mathrm{CPU} \propto N_{\mathrm{MD}} \cdot N_{\mathrm{iter}} \cdot N_k \tag{61}$$

From these considerations it is argued to develop electronic structure methods (cf. Sec. 5) with efficient basis sets to reduce their number N, to develop algorithms to accelerate the convergence (cf. Sect. 2.3) and to employ an efficient $\mathbf{k}$-point integration scheme (cf. Sec. 2.2). While Eq. (61) shows the scaling of the CPU time due to the loop structure the total scaling also

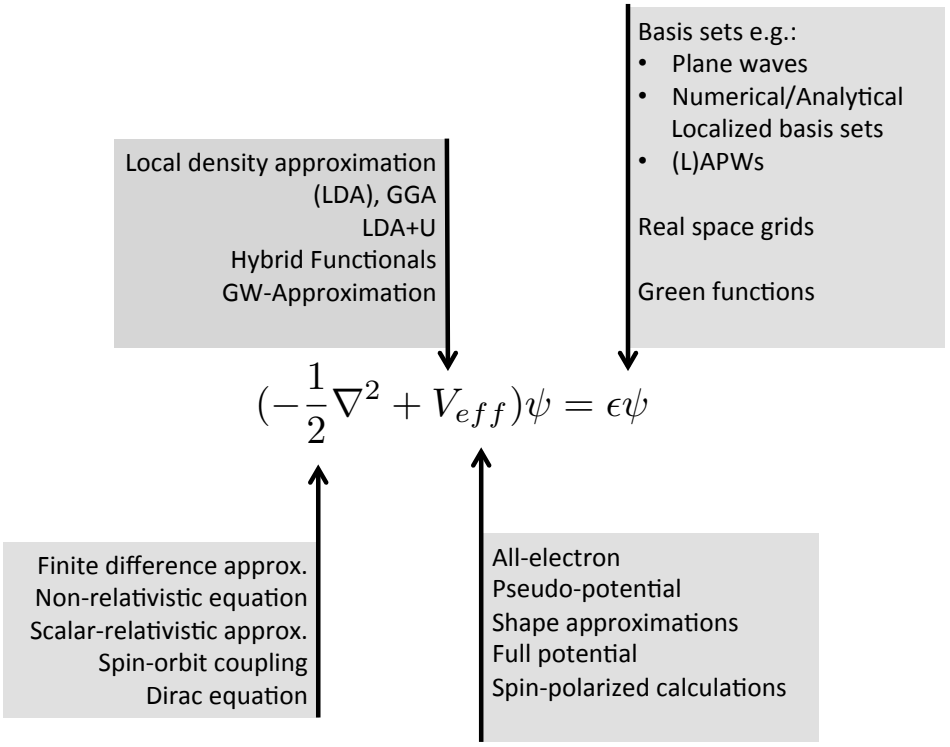


Fig. 10: *Rough diagram indicating the different approximations and numerical schemes applied in DFT codes. The combination of the choices made for each aspect determine the physical properties available and the class of phenomena one can study with the code in question.*

strongly depends on the scaling of the single diagonalization. If a standard dense matrix solver is used this part scales as

$$\text{CPU} \propto N^3.$$

A lot of effort is devoted into finding and tuning the most appropriate eigenvalue solvers to overcome this scaling behavior. In particular, in situations in which only very few eigenvectors of a large sparse matrix have to be computed substantially better scaling can be achieved. For more details on eigenvalue solvers see the contribution D3 by E. Di Napoli.

5.2 DFT on large parallel computers

The key requirement to use DFT on modern computers is the efficient and massive parallelization of the code. Unfortunately, looking at the general self-consistency cycle, one usually identifies several points at which significant CPU time is spend. In most cases the eigenvalue problem is most significant, but also other parts like the construction of the Hamiltonian, the generation of the potential or of the charge can be time consuming as well. The exact ratio of time spend in these parts additionally strongly depends on the method used, i.e. the basis set and other approximations, on the size of the system, its symmetry and so on. Nevertheless, one usually finds some common points in DFT codes that can be exploited for parallelization. First, a single self-consistency cycle can contain loops over relatively independent calculations, such as multiple spins or $\mathbf{k}$-points. These lead to separate eigenvalue problems, which can be dis-

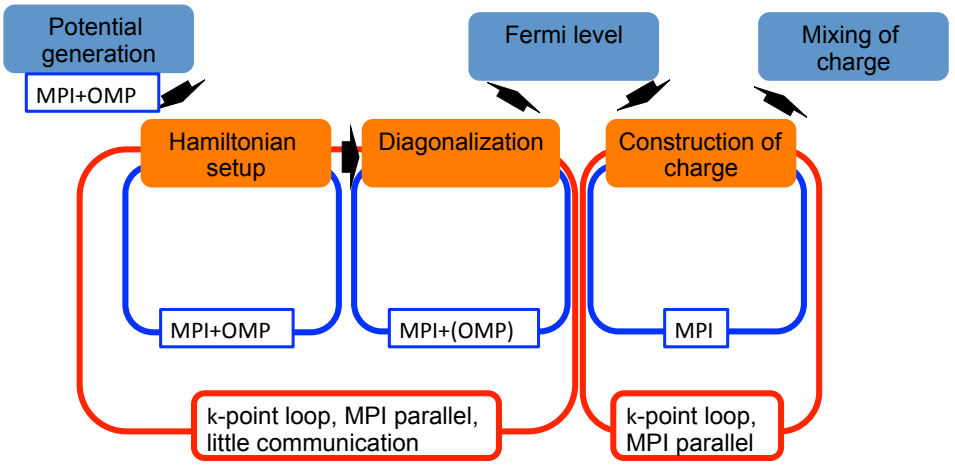


Fig. 11: *Schematics of the parallelization of the FLEUR code (http:www.flapw.de). The different levels of parallelization and the different frameworks used are indicated.*

tributed efficiently. After the eigenvalue problems are solved, only the eigenvalues are needed for the determination of the Fermi-level and then the charge can be calculated from the distributed wavefunctions. This parallelization introduces very little communication and thus is most frequently applied.

However, there are only (maximally) two spins and, as the Brillouin zone shrinks with increasing system size, one usually has to deal with very few $\mathbf{k}$-points in larger calculations. Hence, this kind of parallelization might enable the use up to approximately hundred processors but is insufficient if one wants to use > 1000 cores. Therefore, the massive parallelization of DFT codes requires a multi-level approach(cf. Fig. 11), in which one first parallelizes the easy loops and then additionally distributes the resulting eigenvalue problem further by parallelizing the eigenvalue problem, the charge density generation or other steps. While the parallelization of the outer loops is usually done by a distributed memory approach via the MPI library, the innermost levels of parallelization can also be performed using a shared memory approach using OpenMP or GPU based ideas. In this step of additional parallelization of the eigenvalue problem methods which are based on a real space basis set, i.e. either using localized basis functions or directly a real-space grid can employ the domain decomposition scheme. As the potential term is local in DFT, only little communication with next nearest neighbor processors in a 3D processor grid is needed to evaluate the Hamiltonian matrix elements. Together with iterative solvers for the eigenvalue scheme one obtains algorithms with scale extremely efficiently with increasing processor number as long as this nearest neighbor communication is fast. Fortunately, the network topology of many modern supercomputers is tuned for exactly such kind of communication. The biggest remaining obstacle in such massively parallelized codes frequently is the required orthogonalization of the wavefunctions which requires to exchange information on the wavefunctions which otherwise could efficiently be distributed on disjunct sets of processors. Some example of such DFT calculations of massively parallel machines are given in the contribution D6 of R. Zeller.

Acknowledgment

Parts of this manuscript are taken from the contribution A8 by S. Blügel in the IFF spring school 2006.

References

[1] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).

[2] J. Stoer, *Numerische Mathematik 1*, (Springer-Verlag, Berlin, 1994).

[3] D. J. Chadi and Marvin L. Cohen, Phys. Rev. B **8**, 5747 (1973).

[4] S. L. Cunningham, Phys. Rev. B **10**, 4988 (1974).

[5] O. Jepson and O. K. Andersen, Solid State Commun. **9**, 1763 (1971).

[6] G. Lehmann and M. Taut, phys. stat. sol. (b) **54**, 469 (1972).

[7] P. E. Blöchl, O. Jepsen, and O. K. Andersen, Phys. Rev. B **49**, 16223 (1994).

[8] S. Blügel, (Forschungszentrum Jülich, Jül. Report 2197, 1988).

[9] G. Kerker, Phys. Rev. B **23**, 3082(1981)

[10] P.M. Anglade and X. Gonze, Phys. Rev. B **78**, 045126(2008).

[11] S. Blügel, (PhD thesis, RWTH Aachen, 1988).

[12] D. Wortmann, H. Ishida, and S. Blügel, Phys. Rev. B **65**, 165103 (2002).

[13] D. Wortmann, H. Ishida, and S. Blügel, Phys. Rev. B **65**, 165103 (2002).

[14] E. E. Krasovskii, Phys. Rev. B **70**, 245322 (2004).

[15] P.E. Blöchl, C.J. Först, and J. Schimpl, Bulletin of Materials Science **26**, 33 (2003).

[16] O. K. Andersen, Phys. Rev. B **12**, 3060 (1975).

[17] A. R. Williams, J. Kübler, and C. D. Gelatt Jr., Phys. Rev. B **19**, 6094 (1979).

[18] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).

[19] M. Weinert, E. Wimmer, and A. J. Freeman, Phys. Rev. B **26**, 4571 (1982).

[20] J. C. Slater, Phys. Rev. **51**, 846 (1937).

[21] J. C. Slater, Advances in Quantum Chemistry **1**, 35 (1964).

[22] T. Loucks, *Augmented Plane Wave Method*, (Benjamin, New York, 1967).

[23] H. Bross, G. Bohn, G. Meister, W. Schubo, and H. Stohr, Phys. Rev. B **2**, 3098 (1970).

[24] D. D. Koelling and G. O. Arbman, J. Phys. F (Metal Phys.) **5**, 2041 (1975).

[25] R. V. Kasowski, Phys. Rev. B **8**, 1378 (1973).

[26] P. M. Marcus, Int. J. Quantum Chem. Suppl. **1**, 567 (1967).

[27] M. Weinert, *Solution of Poisson's equation: beyond Ewald-type methods*, J. Math. Phys. **22**, 2433 (1981).

[28] D. R. Hamann, Phys. Rev. Lett. **42**, 662 (1979).

[29] H. J. F. Jansen and A. J. Freeman, Phys. Rev. B **30**, 561 (1984).

[30] S.-H. Wei, H. Krakauer, and M. Weinert, Phys. Rev. B **32**, 7792 (1985).

[31] S.-H. Wei and H. Krakauer, Phys. Rev. Lett. **55**, 1200 (1985).

[32] L. F. Mattheiss and D. R. Hamann, Phys. Rev. B **33**, 823 (1986).

[33] P. Blaha, K. Schwarz, P. Sorantin and S.B. Trickey, Comp. Phys. Commun. **59**, 399 (1990).

[34] see `http://www.flapw.de`

[35] H. Krakauer, M. Posternak and A. J. Freeman, Phys. Rev. B **19**, 1706 (1979).

[36] D. R. Hamann, L. F. Mattheiss and H. S. Greenside, Phys. Rev. B **24**, 6151 (1981).

[37] E. Wimmer, H. Krakauer and A. J. Freeman, Adv. Electronics Electron Phys. **65**, 337 (1985).

[38] In the scalar relativistic approximation (SRA) [49, 39, 40, 41] of the full relativistic Kohn–Sham equations the mass-velocity and Darwin terms are included to all orders of $(1/c^2)^n$, where c is the velocity of light, but the spin–orbit interaction is systematically omitted. In this approach the angular momentum quantum number ℓ and the spin quantum number σ remain good quantum numbers.

[39] T. Takeda, Z. Physik B **43**, 32 (1978).

[40] J. H. Wood and A. M. Boring, Phys. Rev. B **18**, 2701 (1978).

[41] H. Gollisch and L. Fritsche, phys. stat. sol. (b) **86**, 145 (1978).

[42] C. Li, A. J. Freeman, H. J. F. Jansen, and C. L. Fu, Phys. Rev. B **42**, 5433 (1990).

[43] J. M. Soler and A. R. Williams, Phys. Rev. B **42**, 9728 (1990).

[44] R. Yu, D. Singh, and H. Krakauer, Phys. Rev. B **43**, 6411 (1991).

[45] D. Singh, Phys. Rev. B **43**, 6388 (1991).

[46] E. Sjöstedt, L. Nordström, D. Singh, Solid State Commun. **114**, 15 (2004).

[47] T. Takeda and J. Kübler, J. Phys. F **9**, 661 (1979).

[48] S. Goedecker, Phys. Rev. B **47**, 9881 (1993).

[49] D. D. Koelling and B. N. Harmon, J. Phys. C (Solid State Physics) **10**, 3107 (1977).

[50] W. Ning, C. Kailai, and W. Dingsheng, Phys. Rev. Lett. **56**, 2759 (1986).

[51] J. Korringa, Physica **13**, 392 (1947).

[52] W. Kohn and N. Rostoker, Phys. Rev. **94**, 1111 (1954).

[53] T. H. Dupree, Ann. Phys. (N. Y.), **15**, 63 (1961); J. L. Beeby, Proc. Roy. Soc. London Ser. A**302**, 113 (1967); G. J. Morgan, Proc. Phys. Soc. **89**, 365 (1966).

[54] For a review of the KKR method, see N. Papanikolaou, R. Zeller, and P. H. Dederichs, J. Phys.: Condens. Matter **14**, 2799 (2002).

[55] R. Zeller and P. H. Dederichs, Phys. Rev. Lett. **42**, 1713 (1979).

[56] R. Zeller, P. H. Dederichs, B. Újfalussy, L. Szunyogh, and P. Weinberger, Phys. Rev. B **52**, 8807 (1995); K. Wildberger, R. Zeller, and P. H. Dederichs Phys. Rev. B **55**, 10074 (1997).

[57] R. Zeller, J. Phys. C **20**, 2347 (1987).

[58] J.E. Inglesfield, J. Phys. C (Solid State Physics) **14**, 3795 (1981).

D 5 GPUs and Other Non-Standard Hardware[1]

A. V. Adinetz
Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH

Contents

[1]Lecture Notes of the 45^{th} IFF Spring School "Computing Solids - Models, ab initio methods and supercom-

1 Introduction

Graphics Processing Units (GPUs) are throughput-optimized processors originally used to accelerate 3D graphics. Since 2000, they have been increasingly popular for solving computationally intensive problems not directly related to computer graphics, an approach named *General-Purpose GPUs (GPGPUs)* . Using GPUs can be beneficial due to their high resource of parallelism and higher compute power and memory bandwidth available when compared to CPUs. For many real applications, computer systems using GPUs provide better energy efficiency, i.e. performance per watt, and smaller cost of performance, i.e. performance per dollar spent, compared to CPU-only systems.

GPGPU is currently an established technology, and a large and thriving ecosystem is built around GPUs. A lot of GPU-related research is performed, with GPU papers always being a part of conferences on computer and comptuational science, and supercomputing. Courses on GPU programming are taught at universities, and there is a lot of GPU-enabled applications, both open-source and commercial. GPUs are used in fields ranging from video and image processing to astrophycis to seismic modelling to quantum physics. Servers and computer clusters with built-in computational GPUs are commercially available. Twice, in November 2010 and November 2012, has a GPU-based system taken #1 spot of Top 500 list of the most powerful supercomputers. The latest November 2013 edition of the list includes 40 GPU-based systems [1] . And while being used in top supercomputers, GPUs are also widely available in desktop PCs and notebooks, where they allow to reach levels of performance previously achievable only with computing clusters. Due wide availability of GPUs, GPGPU is actually very democratic, as GPU programming does not require any expensive specialized hardware. And of course, this "democraticity" has played a significant role in wide adoption of GPGPU computing.

This lecture attempts to describe the main principles behind GPU architectures as well as demonstrate basics of GPU programming. It focuses mainly on NVidia GPUs, since they are the ones most widely used for computation; however, same architectural principles also apply to AMD GPUs, and to a lesser extents, Intel Xeon Phi accelerators. These lecture notes are organized as follows. We first talk about architecture of modern GPUs in section 2. We then discuss CUDA as the main tool currently used to program GPUs in section 3, and present an example of doing matrix-vector multiplication using GPUs. We discuss other GPU programming tools, libraries and packages in section 4, and make concluding remarks in section 5.

2 GPU Architecture

2.1 Brief History of GPUs

First 3D GPUs appeared to generate 2D images from polygonal models of 3D scenes, or perform *rendering*, in real-time, which is a computationally intensive task. Each vertex should first be projected on the screen, which is done by multiplying its coordinates by projection matrix. After that, the projected triangles are *rasterized*, i.e. converted in the *fragments* or pixels they cover, with invisible pixels culled. Finally, color is computed for each pixel using texture interpolation. One of the ways to get the necessary performance was to use specialized hardware; this, for instance, enabled to perform rasterization very fast. Another possibility of acceleration comes from the fact that rendering actually involves lots of similar operations which do

independent processing of small data elements. Such operations can therefore be performed *in parallel*, and performance of GPUs could simply be increased by using more processing units. Thus GPUs got their resource of parallelism, which is crucial for GPU computing.

Soon it turned out that fixed pipeline was not flexible enough. The transform and texturing stages were replaced in 2001 by *vertex shader* and *fragment shader*, respectively, small programs that enabled usage of developer-defined logic. First shaders allowed only short sequences of only arithmetic operations. Even conditionals had to be emulated through linear interpolation. Programmability improved quickly with time; in 2004, GeForce 6x already enabled several nested loops and conditionals inside the shader.

In 2006, graphics hardware had two processing engines, one for vertex and pixel shaders, with developers often having to balance workload between the two. Also, the third programmable stage, so-called *tesselation shader* appeared. Obviously, there was a need to unify the hardware used to execute different shaders. And if that, why not reverse it: instead of making a special-purpose GPU programmable, make a massively parallel general-purpose processor and implement graphics functions on top of it? This was achieved in 2006 with GeForce GTX 8800 and in 2007 with the release of Compute Unified Device Architecture, CUDA, by NVidia. While it still contained special-purpose hardware for triangle rasterization and texture interpolation, it was a general-purpose parallel processor which could be programmed using CUDA C, a dialect of C language. Design of modern GPUs, though changed a little, still follows the overall pattern of GTX 8800.

GPUs were used to accelerate operations other than 3D rendering even before they became programmable. Programmability available with vertex, and especially pixel shaders, enabled to use GPUs to accelerate image and video processing. For more complex programs, lack of control flow was compensated with invoking the kernel multiple times. With GeForce FX, true single precision became available on the GPU for the first time, and later GPUs added better support for complex control flow, which enabled implementing more complex processing. At this stage, GPUs were already tapped in for such applications as molecular modeling [2], but their processing power was mostly used for visual effects. Compute applications of the era mostly worked with GPUs through 3D APIs, such as OpenGL [3] or Direct3D, GPU programs had to be written in specialized *shader languages*, such as GLSL, and as a result, GPU programming was very cumbersome.

GPU computing changed with the advent of CUDA [4] in 2007. A simple API enabled interaction with GPUs, and GPU programs could be written essentially in C, with most language features available. As a result, GPUs were adopted by a wide range of applications in different domains, with CUDA C being the implementation language of choice. Many packages and libraries, both commercial and open-source, now provide GPU support [5]. New applications demanded new features, which were provided in subsequent CUDA versions; these included atomic operations, accessing RAM of the host from GPU, GPU debugging support, interacting with GPU asynchronously, and with CUDA 5.0 and latest GPUs, even the ability to launch new GPU programs directly from inside running GPU programs.

Success of CUDA was followed by OpenCL [6]. Unlike CUDA, which is a tool proprietory for NVidia, OpenCL is an open standard supported on a wide variety of devices by various manufacturers. OpenACC [13] provides a still higher level of abstraction using directives inserted into code, thus providing higher productivity for GPU software development. If maximum flexibility is required, however, CUDA and OpenCL are still needed.

2.2 Overview of GPU Architecture

Fig. 1 demonstrates the overall place of GPU in a computer. GPU is an *accelerator*, a throughput-optimized processor which is separate from the main CPU, and usually placed on a separate card. It is connected to the CPU through a high-speed link, currently PCI-Express (PCI-E). In this context, CPU is called a host processor, or simply *host* . Any data processing done by the GPU should be set up and requested by the host, by copying data in, launching GPU programs, or *kernels* , and then transferring data out; GPU is, in this sense, is a subordinate processor.
GPU has its own RAM, which is separate from the RAM of the host. It is quite large, up to 12 GiB as of end 2013, but is still smaller than host RAM. A GPU can access its RAM faster than CPU: Kepler K20X GPU, for instance, achieves 160 GiB/s real bandwidth, comapred to 40 GiB/s per CPU, or 80 GiB/s for 2-socket server, for a typical high-end Intel Xeon CPU as of 2013. Modern GPUs can access host memory, but real speed of PCI-E link is much smaller: it is around 11 GiB/s for PCI-E Gen. 3. It is therefore crucial that *data processed by the GPU is stored in GPU RAM, in order to achieve maximum performance*.

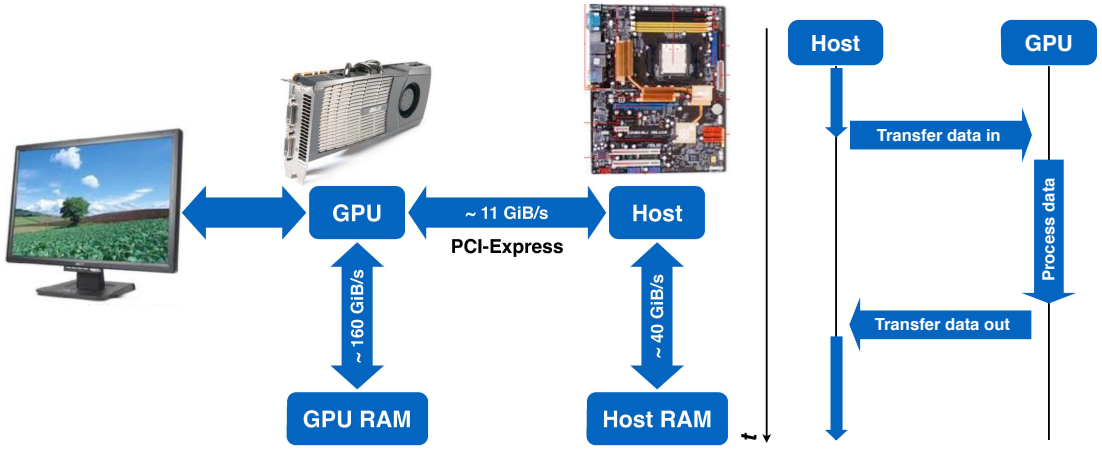


Fig. 1: *Place of a GPU in a computer (left) and typical workflow of GPU processing (right)*

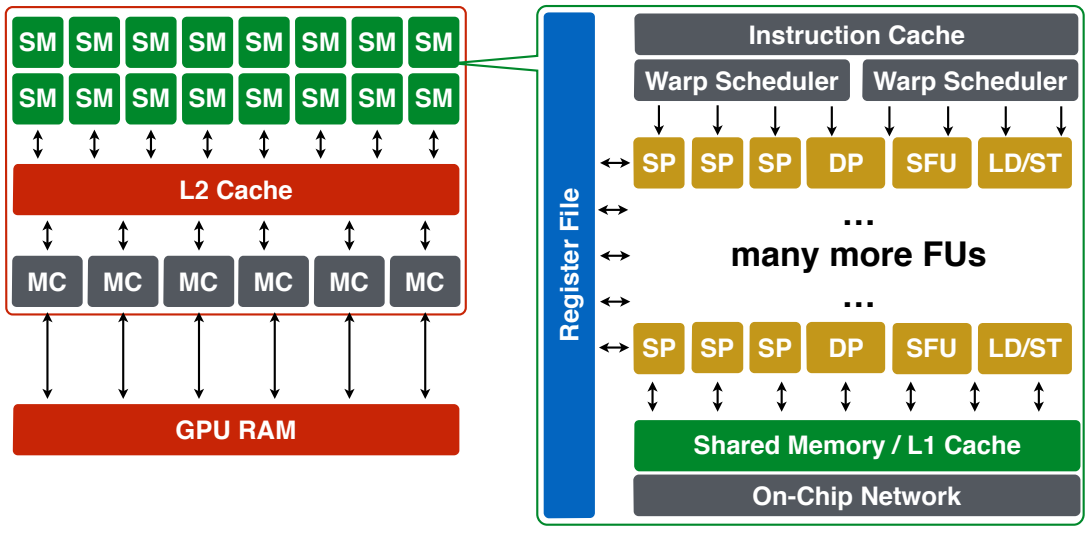


Fig. 2: *Architecture of a GPU chip (left) and a single SM (right).*

Architecture of a GPU chip is depicted in Fig. 2. The GPU itself consists of a number of *streaming multiprocessors (SMs)*, which share a common L2 cache, and a pool of memory controllers (MCs). Each SM contains a lot, often hundreds of, *functional units (FUs)*, which provide the computing power of the GPU. They include FUs for single-precision (SP) and double-precision (DP) computations, special function units (SFUs) for evaluating math functions, and load-store units (LD/ST). Units for doing arithmetic operations are called *Arithmetic-logical units (ALUs)*. An SM also contains schedulers, register file, on-chip *shared memory*, L1 cache, and logic to connect to the rest of the GPU. There is no single understanding of what is called a "GPU core": sometimes, an SM is called a core, and sometimes a single-precision FU inside an SM. Top500 counts an SM as a core.

From the developer perspective, a basic unit of GPU execution is a *GPU thread*. On the level of hardware, the basic unit of execution and scheduling is *warp*, a group of a small number of threads executing synchronously, in a Single Instruction Multiple Data (SIMD) fashion. A warp on an NVidia GPUs consists of 32 threads. There are usually many more warps executing per SM than there are FUs available, with *warp schedulers* used to manage resources. As variables of all threads are stored in registers, an instruction from any thread ready for execution can be scheduled. There are no context switches on GPU; how it works can rather be described as *hardware multithreading*.

Several warps are joined in a *thread block*. Threads of a single block can together use *SM shared memory*, which is present at every SM and can be accessed at the speed of L1 cache. They can also perform *barrier synchronization*, where each thread waits until all other threads arrive at the barrier before continuing execution. Barriers are usually used to synchronize access to shared memory. Barrier synchronization is not possible between threads from different blocks, as it is not even guaranteed that a given pair of thread blocks will be ever executed simultaneously. Warps from a single thread block always share the same SM; however, there may be many thread blocks executing on the same SM, depending on resources available. All thread blocks of a single kernel launched on a GPU, whether already launched or pending, are called a *grid* . Examples of 1D and 2D GPU execution grids are shown in Fig. 3.

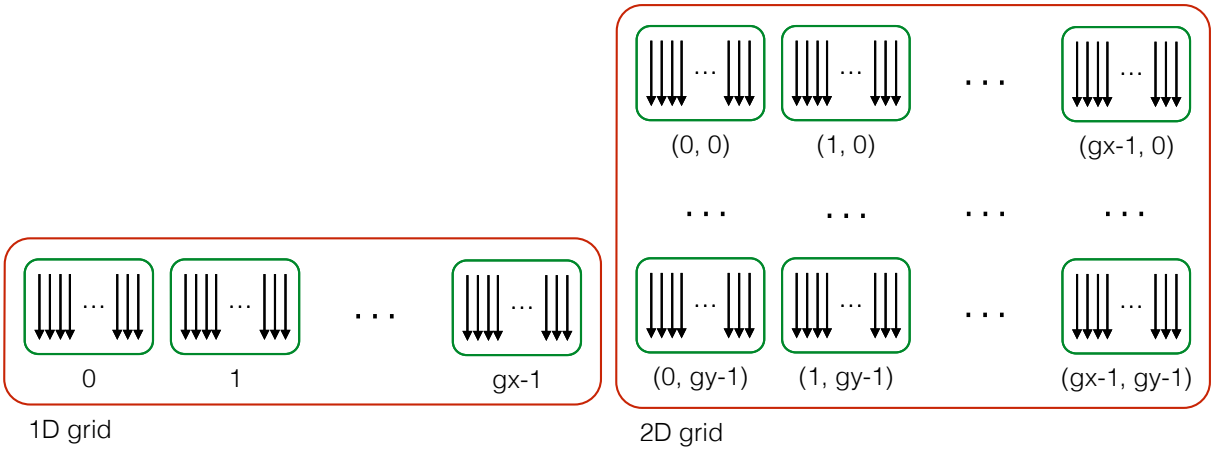


Fig. 3: *Multiple threads and thread blocks executing in 1D (left) and 2D (right) grids on a GPU.*

This architecture outline indicates that *GPUs have very high degrees of parallelism*. As of 2013, high-end CPUs have 8 cores, with two hardware threads executing per core. An average GPU has hundreds of SPs, and high-end ones have thousands. There are 2880 on NVidia Kepler K40, and at least $2\times$ that number of threads is required just to use all hardware available in

case where there are no memory accesses. Still more threads are required to hide memory access latency. As a rule of thumb, there should be $O(10^4)$ threads in a grid running on a GPU just to utilize its resources. This gives an indication of a degree of parallelism required of computational problems to take advantage of computational power of a GPU. On the other hand, GPU threads are cheap, and therefore it is common to launch as many threads as possible to exploit all available parallelism.
Let's consider a simple problem of array addition. A function to do that on a single CPU core will probably look like `add_arrays_cpu`: it contains a loop iterating over all array elements. For multiple CPU cores, there will be multiple threads, but each will execute a similar function for a part of an array. On a GPU, array addition looks completely different: there is a single thread per array element, and the code of a kernel is a code executed by a single thread. One threads per element is actually created during kernel launch; therefore, there is no loop inside a kernel.

```
/* array addition on CPU */
void add_arrays_cpu(float *c, float *a, float *b, int n) {
  for(int i = 0; i < n; i++)
    c[i] = a[i] + b[i];
}

/* array addition on GPU */
__global__ void add_arrays_gpu
(float *c, float *a, float *b, int n) {
  /* get thread id */
  int i = threadIdx.x + blockIdx.x * blockDim.x;
  /* check if thread number is outside range */
  if(i >= n)
    return;
  c[i] = a[i] + b[i];
}
```

Fig. 4: *Adding arrays on a CPU and on a GPU*

2.3 Latency Hiding

To understand latency hiding, let's consider a simple example. Assume a hypothetical CPU, which executes every arithmetic operation in 1 cycle, has 1 core, runs at 2 GHz, has RAM bandwidth 10 GB/s and RAM access latency of 50 cycles. Also assume that a CPU is simple enough that it does not contain any caches or prefetchers (fortunately, the example won't work on a real CPU). This CPU is used to execute addition of floating-point arrays similar to Fig. 4, but it executes slowly, at about 13 MFlop/s [2], and we want to know why. Neither computation nor memory bandwidth are bottlenecks: only 13 MFlop/s of 2 GFlop/s and 52 MB/s of 10 GB/s, respectively. The reality is that the application is *latency-limited*: each memory access

[2]This won't happen on real CPUs, due to hardware prefetchers which predict following memory accesses and retrieve data for them, and caches; however, it is useful as an illustration.

takes long time, and nothing more can take place during that time: computations cannot be performed, and new memory accesses cannot be initiated. Latency-limited behavior is obviously bad, as it leads to severe under-utilization of hardware resources. The problem can be solved using caches; assuming 64B cache lines, a latency of 5 cycles and no hardware prefetchers or speculative execution, up to 82 MFlop/s performance and 328 MB/s memory bandwidth can be achieved. However, this is still a very small fraction of what is achievable.
GPUs solve the latency problem in a different way, using hardware multithreading. Assume the same hypothetical CPU, but which now supports many hardware threads, similar to Fig. 2. When a thread performs a memory access, it is placed in a waiting list, and instructions from other threads are scheduled, when available. With 8 threads per core, 426 MB/s bandwidth is achieved, and with 16 threads, about 853 MB/s is achieved. Further speedup can be achieved by enabling a single thread to issue multiple memory requests. With array addition, each thread issues 2 memory requests, thus incurring memory latency only once. Similarly, if a thread issues a write, it doesn't need to wait for its completion; it may be required that all threads do a barrier at the end of the loop to ensure coherent memory state. Latency is thus incurred only once per iteration, and bandwidth of 1.28 and 2.56 GB/s is achieved for 8 and 16 threads, respectively. Execution timeline will be similar to Fig. 5. If multi-core is added to that, 8 cores with 8 threads each will already be able to saturate entire memory bandwidth, thus achieving best possible performance for the application, which is memory-limited. Of course, all this comes at a cost: memory controllers should be able to handle hundreds of requests simultaneously, and large register files are required to store data of multiple hardware threads on chip. However, this approach allows to achieve *latency hiding*, i.e. turn any latency-limited application into bandwidth-limited, provided sufficient hardware resources.

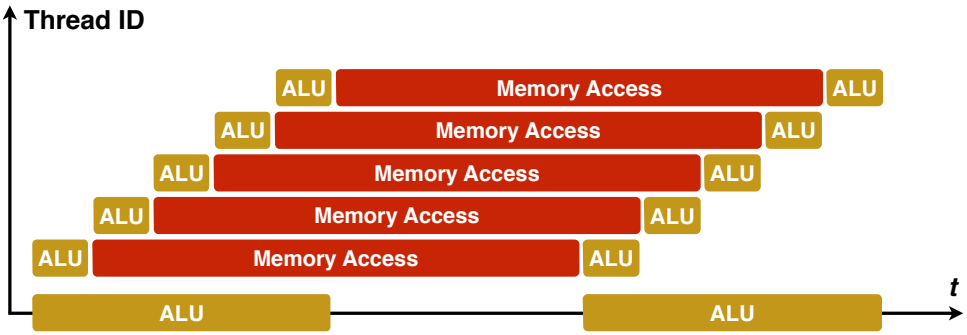


Fig. 5: *Hardware multi-threading used for latency hiding: with many hardware threads, utilization of ALU and memory bandwidth increases*

Latency hiding is used by all modern GPUs; the difference from the above is that warps are used instead of threads, and both bandwidth and latency are an order of magnitude higher. Therefore, the number of threads required for hiding latency is also considerably higher, $O(10^4)$. Another aspect is using hardware multi-threading to hide not only latencies of memory accesses, but also latencies arising from accessing register files and executing arithmetic operations. In this sense, a GPU can be thought of as a *throughput-optimized processor*, i.e. optimized for executing lots of threads to achieve maximum throughput rather than for fastest execution of a single thread. This is the guiding principle of GPU design; therefore, GPUs have lots of very simple functional units, no advanced control features such as branch prediction or speculative execution, and smaller caches. Because of all these features, it is possible to run GPUs at lower frequencies and achieve superior performance and energy efficiency with parallel applications. CPUs, on

the contrary, are *latency-optimized processors*: they execute a single thread considerably faster than a GPU, but are slower on applications with large resource of parallelism.
It is therefore crucial to have enough GPU threads available for execution to hide latencies. *Occupancy* is the ratio of warps executing per SM to the maximum number of warps possible on SM; low occupancy, usually indicates performance problems. Low occupancy is most often caused by excessive use of resources by the kernel, such as registers and shared memory. Note that as all local variables of a GPU thread are stored on registers, as a rule, complex kernels will use more registers and usually have lower occupancy value. Also note that the problem can be aggravated by too large a thread block size: as whole blocks are scheduled on SMs, resource usage can become better if thread block is made smaller. Generally, thread block size of 128–256 is a good starting point for experimentation.

2.4 Other GPU Architecture Principles

As mentioned above, threads in a GPU warp execute synchrnously, in SIMD fashion. But what happens if the threads encounter a conditional construct? If condition evaluates the same for all threads, just one branch is executed. But what if condition evaluates differently, and some threads take **`if`** branch, while the other take **`else`** branch? What actually happens is depicted in Fig. 6: all threads execute *both branches* sequentially. First, all threads take the **`if`** branch. The *active threads*, i.e. ones for which the condition evaluated to true, actually execute as usual, while *inactive threads*, for which the condition evaluated to false, are *masked*: they still execute instructions, but their results are not written to registers. When the **`if`** branch is finished, the same repeats for the **`else`** branch, with active threads becoming inactive, and vice versa. Note that only threads in a warp execute synchronously: when the condition is different across warps in the same thread block, but is constant within the warp, each warp executes only one branch. Outside of warp, a GPU is thus a true Multiple-Instruction Multiple-Data architecture (MIMD).

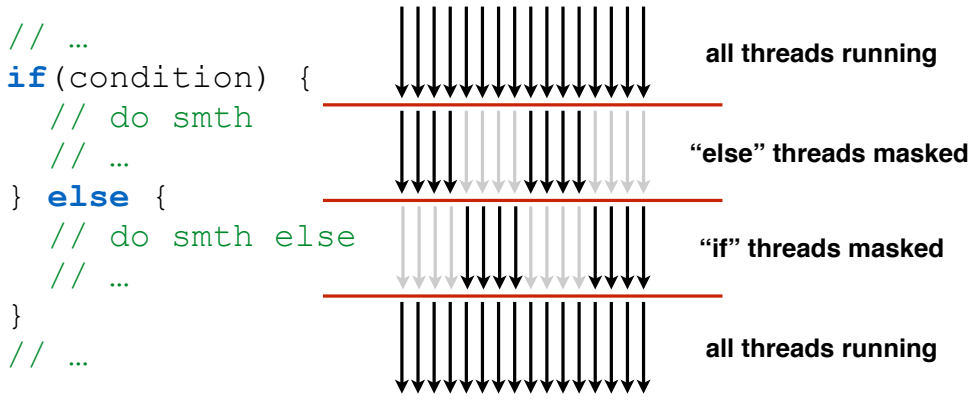


Fig. 6: *Branch divergence of a threads in a GPU warp*

The condition when threads in the same warp take different execution paths is called *branch divergence*, and impacts performance. Imagine that both **`if`** and **`else`** branches contain some computatinoally intensive parts, which have about the same execution cost. Then performance of this code region drop $2\times$ when just a single thread in a warp chooses a different path. With nested conditionals, performance drop can be as high as the number of threads in a warp, i.e. up to $32\times$ on modern NVidia GPUs. To avoid performance impact of branch divergence, it may be required to re-order input or split a complex kernel into simpler parts. Note that for

simple conditionals, such as `a = b > 0.5 ? c : d`, are translated into *conditional move* instructions, and do not incur performance overhead.

Memory accesses from the threads in a same warp are also initiated synchronously. They are not served a single word at a time; instead, memory requests are grouped into *transactions* based on which *memory regions* they access, and executed in units of transactions. A single memory region usually has $N = 2^k$ bytes and aligned to N bytes boundary. On modern GPUs, N varies from 32 to 128, based on whether L1 caching is used; this corresponds to from 8 to 32 4-byte values of type **`int`** or **`float`**. Grouping memory accesses into transactions is called *coalescing*. Using less words than retrieved leads to underutilization of memory bandwidth, and can thus reduce performance. The ratio of words used to words retrieved is called *degree of coalescing*, and corresponds to the degree of bandwidth utilization. Note that it can exceed 1, e.g. when all threads in a warp access the same word. When the ratio is high, accesses are said to be *coalesced*, and when it is $>= 1$, they are *fully coalesced*; when it is low, accesses are said to be poorly coalesced. Several examples of coalesced and uncoalesced access patterns are illustrated on Fig. 7.

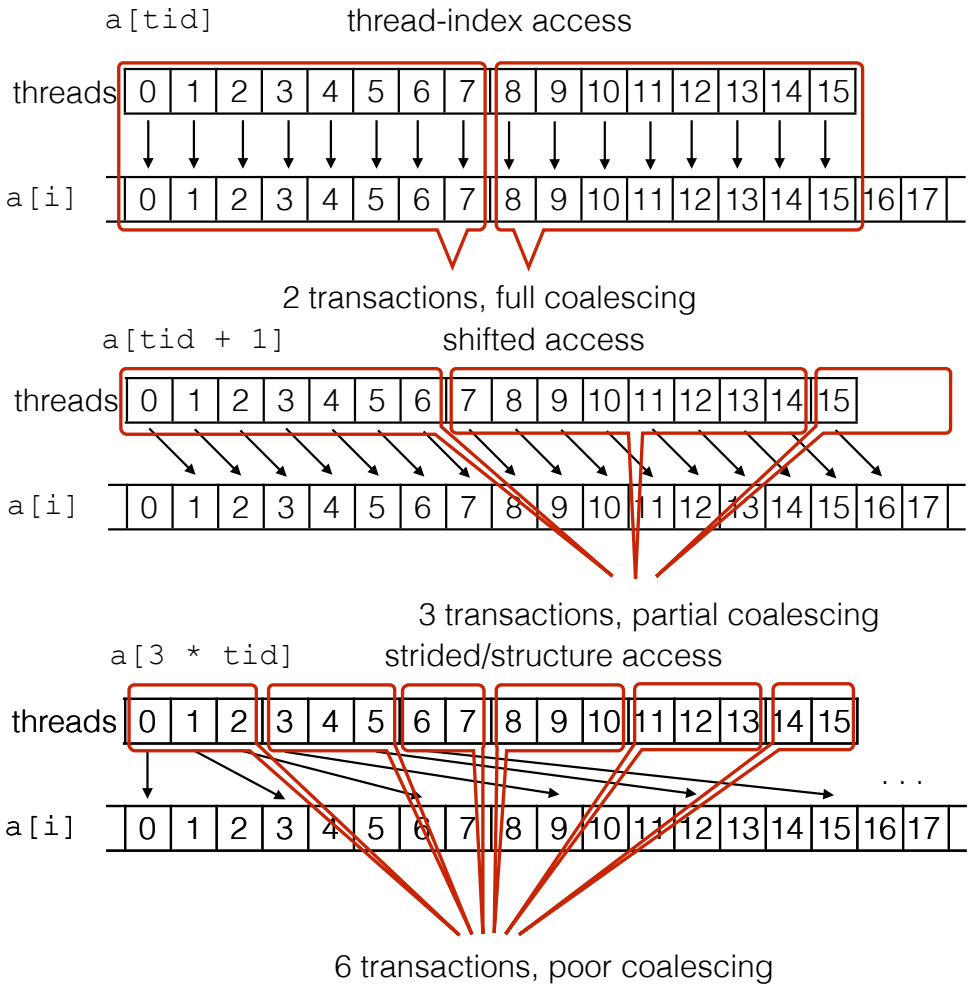


Fig. 7: *Different good and bad coalescing patterns*

Unlike branch divergence, poor coalescing is much easier to encounter in practice, as it is exhibited by a number of common access patterns. One example of this are scattered indirect accesses, or data-dependent accesses, like `c[i] = a[b[i]]`; degree of coalescing is actually

dependent on contents of array `b`. Similarly, poor coalescing can result if threads in the same warp access data using different pointers, e.g. when traversing tree or graph structures. Yet another poor coalescing pattern is accessing fields of a C struct in an *array of structures (AoS)*; it is usually better to have one array per structure field, or a *structure of arrays (SoA)*. Similarly, accesses to different rows of an array stored in row-major order, i.e. when neighboring elements of the same row are stored at adjacent indices, exhibits poor coalescing. Poor coalescing can also have a more devastating effects on performance; while several nested conditionals are required to cause significant branch divergence, just one statement with bad coalescing is enough to cut memory bandwidth by $16\times$. Profiling tools, described in more detail in section 4.5 can help with detecting issues such as branch divergence and bad coalescing.

2.5 Requirements for Applications to be Portable to GPUs

Features of GPU architecture described above lead to requirements to applications, or parts of applications, that can be efficiently executed on a GPU:

- **High degree of parallelism**, i.e. lots of independent operations, is absolutely required to get good GPU performance;
- **Working set must fit into GPU RAM**, so that high-bandwidth RAM can be used.

These requirements may seem hard to fulfill. However, the necessary degree of parallelism has been found in many real applications. The reasons of this are two-fold: first, nature actually "runs in parallel", and thus any simulation of it should exhibit high degree fo parallelism. Second, "Big Data" applications process lots of data, and the volumes of data always grow; processing data elements independently gives high degree of parallelism. Moreover, parallelism, though in a lower degree, is required to make use of traditional supercomputers as well. In fact, it turns out that it is actually hard to find an application which does not exhibit a good degree of parallelism, at least when processing large quantities of data. With regards to the working set, it is often possible to split it into smaller parts that fit into GPU memory.
In addition, it is important that the application exhibits **good memory coalescing**, to make good use of GPU RAM bandwidth and **coherent execution** so that there is no branch divergence. An application is often needed to be modified to achieve this. Additionally, an application should usually exhibit **high computational intensity**, to be able to make good use of GPU compute power. However, even if the latter is not the case, the application can still be accelerated by the GPU, e.g. because of higher memory bandwidth.

3 GPU Programming with CUDA

3.1 CUDA Overview

CUDA, which stands for Compute Unified Device Architecture, is a set of tools to program NVidia GPUs. It consists of the following components:

CUDA Architecture the architecture of NVidia GPUs used to run CUDA programs;

CUDA C/C++ a set of extension to C++ language to enable GPU programming, and also simplify host-GPU interaction;

CUDA Driver a part of NVidia GPU driver which supports GPU on OS levels, and also manages GPU resources;

CUDA Runtime a userspace library which provides API for applications to interact with GPUs, e.g. for copying data or launching kernels;

CUDA Toolchain which is a set of compilation, debugging and profiling tools to support GPU programming.

When talking about evolution of CUDA, is it important to distinguish two different concepts of "version". *CUDA version* is the version of CUDA software, i.e. CUDA C++, compiler and API, while *compute capability (CC)* refers to the features supported by the hardware. In addition, CC affects such aspects as memory coalescing or the number of registers available to the application. To use a specific feature in your program, it must be supported by both CUDA version and CC, and the application should be compiled for a specific CC. If a certain feature is not supported by hardware CC, it cannot be used even if CUDA version supports it. For instance, if a Fermi GPU (CC 2.0) and CUDA 5.0, which supports *dynamic parallelism*, are installed in the machine, then dynamic parallelism still cannot be used, as it requires a device with CC $>=$ 3.5.

3.2 CUDA C++ Langauge Extensions

CUDA introduces 3 additional function qualifiers: **__host__**, **__device__** and **__global__**. Possible combinatons of the qualifiers are listed in table 1. If no qualifier is used, default qualifier is **__host__**, so that CUDA compiler is able to compile existing host code with no change. If both **__host__** and **__device__** qualifiers are specified, two versions of the function are generated, one for host, and one for GPU, so that codebase can be shared between different processors.

Qualifier	Runs on	Called from	Is a kernel?
__host__	host	host	no
__device__	GPU	GPU	no
__host__ __device__	host, GPU	host, GPU	no
__global__	GPU	host GPU (only with dynamic parallelism, CC 3.5+)	yes

Table 1: *CUDA function qualifiers*

__global__ denotes a kernel function, or kernel, that specifies code for a single thread executed on the GPU. A kernel function cannot be called directly as other functions. A special construct, called *kernel launch construct* should be used to start a kernel with a configuration of many threads. As there are many threads, a kernel cannot return a value, and should have return type **void**. It is possible to launch 1D, 2D and 3D execution grids, each consisting of 1D, 2D or 3D thread blocks. Datatype `dim3`, which is a vector of 3 ints, is used to specify grid and thread block size. For 1D launches, however, integer types can be used. A simplified form of a kernel launch construct is

`kernel_name<<<Ng, Nb>>>(arg0, ..., argn)`

where `Ng` is a grid size, `Nb` is the thread block size, and `arg0, ..., argn` are kernel arguments which are passed to each kernel thread. If `Ng` is $g = gx \times gy \times gz$, and `Nb` is $b = bx \times by \times bz$, then g thread blocks, each with b threads are queued for launch. *The kernel launch construct returns immediately*; it does not wait for the kernel to end or even to start. The number of thread blocks executed simulataneously depends on the kernel resource usage, available hardware resources and other kernels currently executing. Fig. 8 shows two equivalent ways to launch the array addition kernel: one with integer sizes, and one with explicit `dim3` variables. Note that the total number of threads launched must be a multiple of thread block size, and thread block size 1 is very bad. Therefore, the number of array elements must be first rounded up to the next multiple of thread block size, and then the extra threads should be prevented from doing anything in the kernel code, as in Fig. 4. For 1D configuration, y and z components are set to 1.

```
// arrays in device memory, their size
float *a, *b, *c;
int n;
// ...

// launch configuration with dim3
int bs = 128;
dim3 block(bs, 1, 1), grid(n / bs + (n % bs ? 1 : 0));
add_arrays<<<grid, block>>>(c, a, b, n);

// launch config with integers
int bs = 128, ng = n / bs + (n % bs ? 1 : 0);
add_arrays<<<bs, ng>>>(c, a, b, n);
```

Fig. 8: *Two launch configurations for array addition kernel*

Different GPU kernel threads usually process different elements of data: in case of array addition, they add array elements of different indices. For a thread to know which elements it has to add, the thread needs to find out its identifier. In CUDA, thread id is separated into *thread block id* and *thread id inside block*. Both of these, as well as the execution configuration, can be discovered through built-in variables, which are listed in the table 2; `uint3` type is similar to `dim3`.

Variable name	Variable type	Meaning
`gridDim`	`dim3`	grid size, in thread blocks
`blockIdx`	`uint3`	current thread block index
`blockDim`	`dim3`	thread block size
`threadIdx`	`uint3`	index of current thread inside block
`warpSize`	`uint`	warp size (currently 32)

Table 2: *CUDA built-in variables*

With variables defined above, for 1D configuration, global thread index can be computed as

```
int i = threadIdx.x + blockIdx.x * blockDim.x;
```

The indices always contain 3 components, `x`, `y` and `z`. For 1D configurations, `y` and `z` are equal to `1`. For 2D configurations, only `z` is equal to `1`, and thus, both `x` and `y` components are needed for a global index.
Variable qualifiers are listed in table 3; they apply to variable declarations. **`__constant__`** and **`__device__`** qualifiers are applied to global variables, while **`__shared__`** qualifier, except for one case, is used with local variables. Constant memory variables are stored in global memory, are read-only on GPU, and use highly efficient constant caches on device. Constant memory can be modified by the host between kernel launches, and all data there cannot exceed 64 KiB in size. Note that qualifiers are only applied to variable declarations; pointers are left unqualified.

Qualifier	GPU Memory	Variable space
`__device__`	global	global
`__constant__`	constant	global
`__shared__`	shared on-chip	local

Table 3: *Variable qualifiers in CUDA 5.5*

While most language features are supported in GPU code, most of the standard library functions are not. And of those features that are supported, some are inefficient. Note that CUDA provides its separate set of math functions, which closely follows, but do not directly corresponds to, functions declared in `"math.h"` header file. Also note that C type declarations and **`#define`** definitions, unlike functions, can be freely used in GPU code, e.g. to declare variables. What follows is a list of features that either are not supported or can lead to severe performance penalties:

- input/output and system calls;
- dynamic memory allocation, i.e. `malloc/free` (thought supported, incurs a much higher performance penalty compared to CPU);
- standard C library functions (printf to stdout is supported, as are most math functions);
- recursion, virtual functions and call by function pointer (lead to severe performance penalty, as functions on GPU are normally inlined; also, register usage cannot be estimated, which severely limits the number of threads launched);
- C++ exceptions and runtime type information (RTTI);
- bitfields in structures;
- **`static`** variables.

3.3 CUDA Runtime API

All CUDA Runtime functions return error code of type `cudaError_t`. Return value of `cudaSuccess` indicates success, and other values indicate failure. CUDA API functions are usually called from the host. However, with dynamic parallelism, it is possible to call some of

those functions on the GPU. What follows is a list of most important CUDA functions which enable writing simple CUDA applications. Refer to the NVidia documentation website [17] for the complete documentation on CUDA Runtime API.
CUDA functions for working with memory are designed based on similar functions in the standard C library. Memory management is performed using the following functions:

```
cudaError_t cudaMalloc(void **pptr, size_t nbytes)
cudaError_t cudaFree(void *ptr);
```

`cudaMalloc` allocates `nbytes` bytes of memory, and saves the pointer to it in `*ptr`. As the return value is the error code, the poiner cannot be returned directly. The pointer returned has the same size as pointer on host, and be freely passed around, and to device functions. However, dereferencing it in host code will lead to undefined behavior, most likely a segmentation fault. To free device memory, call `cudaFree` function.

Data has to be copied from host to GPU before processing, and copied back after. In addition, some data will be on GPU from the start; it is often a good idea to have it initialized. This is achieved using the following functions:

```
cudaError_t cudaMemcpy
  (void *dst, void *src, size_t nbytes, cudaMemcpyKind kind);
cudaError_t cudaMemset(void *p, int value, size_t nbytes);
```

`cudaMemcpy` is used to copy `nbytes` bytes from address `src` to address `dst`. `kind` indicates whether the addresses reside on host or on device; it can be a constant of form `cudaMemcpyXxxToYyy`, where `Xxx` and `Yyy` can be either `Host` or `Device`, for a total of 4 possible combinations. In addition, when *unified virtual addressing (UVA)* is enabled, the runtime can determine whether the address belongs to a host or a device based on the poiner value; in this case, `cudaMemcpyDefault` can also be specified.
Normally, host data resides in *paged memory*, i.e. memory that can be swapped to disk. When a CPU references a page that has been swapped to disk, an interrupt occurs, which is handled by the OS which brings the page back into RAM. However, devices accessing host RAM do not support such logic. Therefore, before the data is copied to the GPU, it has to be copied to *pinned memory*, which is not swapped to disk, and then to the GPU. Such double copying can lead to performance degradation. It is therefore advisable that the data copied to the GPU is allocated in pinned memory. This can be done using `cudaMallocHost()` function. Such memory can be freed using `cudaFreeHost()` function; their signatures are similar to `cudaMalloc()` and `cudaFree()` respectively. Performance of copying from pinned memory is about $2\times$ higher than when copying from paged memory.

The following functions are available for synchronization with the GPU:

```
cudaError_t cudaDeviceSynchronize(void);
cudaError_t cudaStreamSynchronize(cudaStream_t s);
```

`cudaDeviceSynchronize` function waits for all commands previously issued on the current device to be finished. `cudaStreamSynchronize` waits only for the commands on the specified *stream*. A stream is a queue of command submitted to GPU, such as kernel launches and memory copies, executed sequentially. Stream `0` or `NULL` is the default stream for the GPU; waiting for it also waits for commands queued on other streams (unless they opted out of

it). Note that both functions can return error codes from previous asynchronous commands.

The following functions enable querying devices and other device operations:

```
cudaError_t cudaGetDeviceCount(int *ndevices);
cudaError_t cudaGetDevice(int *idevice);
cudaError_t cudaSetDevice(int idevice);
cudaError_t cudaDeviceReset(void);
```

`cudaGetDeviceCount()` gets the total number of CUDA-capable devices attached to host, which is returned in memory pointed to by `ndevices`. The devices are numbered from `0` to `*ndevices - 1`, and are referred by their number. `cudaGetDevice()` gets the current device, while `cudaSetDevice()` sets the current device to `idevice`. The current device is the one to which all CUDA commands, such as memory copies and kernel launches, are submitted. Setting the current device enables submitting commands to different devices, thus enabling multi-GPU programming. Finally, `cudaDeviceReset()` resets the state of the current device and frees resources associated with it; the device is re-initialized by a subsequent CUDA call. As most programs use the device for their entire duration, and resources are automatically freed upon program exit, this call is rarely used. One of the reasons to call `cudaDeviceReset()` at the end of the program is for NVidia Profiler to work correctly.

Error handling functions are as follows:

```
cudaError_t cudaGetLastError(void);
const char *cudaGetErrorString(cudaError_t err);
```

`cudaGetLastError()` returs the error code of the error that occurred last, or `cudaSuccess` if everything is OK. Error code is normally returned by CUDA calls, so it is usually only called after a kernel launch construct, to check if the kernel has been successfully queued. `cudaGetErrorString()` returns a textual description for an error code, and is useful in handling errors. To catch errors early, it is a good idea to check the return code of every CUDA call; to avoid lots of boilerplate code, such check is usually wrapped in a macro, similar to `cucheck()` defined in Fig. 9. In case an error is encountered, it prints a place what error occured and where, and terminates the application with a failure exit code. This enables to quickly identify the place of code where an error occurred, and take action to correct it.

```
#define cucheck(call)                                          \
  {                                                            \
  cudaError_t cucheck_err = (call);                            \
  if(cucheck_err != cudaSuccess) {                             \
    const char *err_str = cudaGetErrorString(cucheck_err);     \
    fprintf(stderr, "%s (%d): %s in %s\n",                     \
       __FILE__, __LINE__, err_str, #call);                    \
    exit(-1);                                                  \
  }                                                            \
  }
```

Fig. 9: `cucheck` *macro used to check error code of CUDA API calls.*

3.4 "Hello, CUDA World!"

Using a programming language typically starts by writing a "Hello, World!" program for it. For CUDA, however, this makes little sense, as GPU is primarily a computational device not designed to support I/O. Instead, a simple parallel computation is often implemented as a first program; here, we'll use **matrix-vector multiplication** as an example. Fig. 10 illustrates the kernel, while Fig. 11 shows host code that makes use of the kernel. Host code uses `cucheck` macro defined in Fig. 9. We compute $b = Ax$, where A is stored column-first so that accesses are coalesced. Host arrays are allocated and initialized first. Then the GPU arrays are allocated, and host data is copied to them. The kernel is then called, and the data is copied back. As data copy executes in the same stream as the host kernel, it only starts after the kernel finishes execution, and thus copies correct data. Finally, first 10 elements of the resulting vector are computed. Note the use of `h_` and `d_` prefixes for host and device pointers, respectively; it is a useful convention to avoid confusing pointers to host and device memory.

```
__global__ void matvecmul_k
(float *b, const float *a, const float *x, int n) {
  int i = threadIdx.x + blockIdx.x * blockDim.x;
  if(i >= n)
    return;
  float bi = 0;
  for(int j = 0; j < n; j++)
    bi += a[j * n + i] * x[j];
  b[i] = bi;
}  // matvecmul_k
```

Fig. 10: *Matrix-vector multiplication kernel in CUDA*

4 Other GPGPU Programming Tools

4.1 OpenCL

OpenCL [6] stands for "Open Compute Language". It was initially proposed by Apple as an open standard alternative to CUDA in 2008. The Khronos OpenCL Working Group to support the standard was formed the same year. First implementations were released in 2009. As of end 2013, the most recent version of the standard was 2.0, and the most widely supported versions are 1.1 and 1.2 (there's little difference between them).

OpenCL closely follows CUDA in its design. It consists of **OpenCL Runtime**, which is a set of API to interact with the device, and **OpenCL C**, an extension of C language to write programs for the device. Unlike CUDA C/C++, OpenCL C contains only device part of the code; host code has to be written separately in another language, such as C or C++. Therefore, unlike in CUDA, there is no single-source compilation. Moreover, OpenCL C program is typically compiled after the start of the host program; though API supports loading binaries, there is currently no portable way to generate such a binary offline. OpenCL Runtime API is similar to CUDA; however, it is more complicated. The developer has to manually manage OpenCL objects, such as contexts, command queues, buffers and programs, even for very simple programs. Moreover,

```
int main(void) {
  // number of elements in matrix and vector
  int n = 8192;
  size_t vec_sz = n * sizeof(float), mat_sz = n * vec_sz;
  // pointers to data on host and device
  float *h_b = (float *)malloc(vec_sz);
  float *h_a = (float *)malloc(mat_sz);
  float *h_x = (float *)malloc(vec_sz);
  // initialize host data
  for(int i = 0; i < n; i++)
    h_x[i] = random() % 20;
  for(int j = 0; j < n; j++)
    for(int i = 0; i < n; i++)
      h_a[j * n + i] = random() % 20;
  // allocate device data
  float *d_b, *d_a, *d_x;
  cucheck(cudaMalloc((void **)&d_b, vec_sz));
  cucheck(cudaMalloc((void **)&d_a, mat_sz));
  cucheck(cudaMalloc((void **)&d_x, vec_sz));
  // copy data to device
  cucheck(cudaMemcpy(d_x, h_x, vec_sz, cudaMemcpyHostToDevice));
  cucheck(cudaMemcpy(d_a, h_a, mat_sz, cudaMemcpyHostToDevice));
  // call the kernel
  int bs = 128;
  matvecmul_k<<<n / bs + (n % bs ? 1 : 0), bs>>>
    (d_b, d_a, d_x, n);
  cucheck(cudaGetLastError());
  // copy data back
  cucheck(cudaMemcpy(h_b, d_b, vec_sz, cudaMemcpyDeviceToHost));
  // print first 10 elements
  for(int i = 0; i < 10; i++)
  printf("b[%d] = %lf\n", i, (double)h_b[i]);
  // free everything
  cudaFree(d_x); cudaFree(d_a); cudaFree(d_b);
  free(h_x); free(h_a); free(h_b);
  // reset; necessary for profiling
  cucheck(cudaDeviceReset());
}  // main
```

Fig. 11: *CUDA host code for matrix-vector multiplication*

as the host program is written in plain C/C++, there is no counterpart to CUDA kernel launch construct; instead, API functions must be used to set arguments and launch the kernel. As a consequence, a "hello, world!"-style program will be significantly more complicated than with CUDA, and is omitted here for space reasons.

As OpenCL is an open standard, it is supported on a wide range of platforms. As of 2013,

OpenCL programs can run on GPUs from NVidia, AMD and Intel; x86, PowerPC and ARM CPUs; and Intel Xeon Phi accelerators. There is even support for OpenCL compilation for FPGA devices [7]. The downside of standardization is that it is hard to add support for newest hardware features. As a result, new features of NVidia GPUs are much better supported in CUDA than in OpenCL. Also, due to late introduction, most applications have chosen CUDA when opting for GPU support. Therefore, there are much fewer applications using OpenCL than CUDA. OpenCL, along with CUDA, is supported as a target for code generation for PGI and HMPP OpenACC compilers, so it is possible to use OpenCL indirectly in your applications.

4.2 GPU-Supported Libraries and Packages

Since the introduction of CUDA, NVidia heavily focused on building a wide ecosystem of libraries and packages supporting CUDA. Therefore, it is possible, for many applications, to use GPUs without writing any GPU code. A number of libraries come directly with CUDA; while they are not open-source, they can be downloaded for free. These libraries include:

cuBLAS provides BLAS level 1–3 routines for dense linear algebra applications

cuFFT provides fast Fourier transform (FFT) on GPUs, and since CUDA 5.5, also a variant of FFTW library which can be linked with application to provide GPU support

cuSPARSE provides numerous sparse matrix formats and sparse linear algebra routines

cuRAND provides various random number generators (RNG)

NPP NVidia Performance Primitives are routines for image and signal processing

Thrust is a C++ template library designed to accelerate routine data-intensive tasks such as reduction or sorting on GPU; it is described in more details in section 4.3

AMD, the other manufacturer of GPUs, also makes some libraries available with their heterogeneous computing platform [8]. In addition to **Accelerated Parallel Processing (APP) SDK**, **clMath** library is provided to accelerate BLAS rountines and FFT on AMD hardware, and **Bolt** is AMD's counterpart of **Thrust**.

There are a number of other libraries supporting GPUs. **MAGMA** [9] is a large collection of linear algebra routines, **OpenCV** [10] is for machine vision and **QUDA** [11] is a lattice QCD library. Libraries such as MAGMA and OpenCV support multiple backends, including CPUs, OpenCL and CUDA for various GPUs, and other accelerators, and can therefore be used to write portable applications taking advantage of available hardware.

A standard way of programming clusters is *Message Passing Interface (MPI)*, a library for exchanging messages over network. A number of MPI packages are now *CUDA-aware* [12], which means that pointers to GPU memory can be used where only pointers to host memory were allowed before. This simplifies writing applications for computer clusters with GPUs, and also improves performance. Using GPUs is also supported from a number of widely used environments, such as Matlab, LabVIEW, Mathematica and Python.

4.3 Thrust

Thrust is part of CUDA and is a template library inspired by STL: it uses the same concepts, such as functors and iterators. A *functor* is an object which overloads **`operator()`**, and can thus be used in place of a function. Similarly, *iterator* generalizes the concept of pointers. Thrust is useful to accelerate boilerplate data-processing operations on GPU, such as reduction, prefix sums, filtering and sorting. Such operations often occur as parts of real applications, but are tricky to implement and optimize for differnt generations of GPUs. Thrust provides versions optimized for different generations, CCs and input data types, and can automatically select the best version. Moreover, thrust provides a simple wrapper to device pointer, `device_ptr<T>`, which enables seamless using of Thrust in CUDA applications. Fig. 12 shows how computing a sum of a floating-point array can be easily implemented using Thrust.

```
// include necesseary declarations
#include <thrust/device_ptr.h>
#include <thrust/functional.h>
#include <thrust/reduce.h>
using namespace thrust;

// ...

// sum up array elements on GPU
float sum(float *d_a, int n) {
  device_ptr<float> start(d_a), end(d_a + n);
  return reduce(start, end, plus<float>(), 0.0f);
}
```

Fig. 12: *Using Thrust to sum up all array elements*

4.4 Directive-Based Approaches

Programming GPUs with CUDA requires manually separating relevant parts of the code into kernels, and transferring data back and forth. In many cases, the kernel code has to be then manually modified to achieve high performance. Often, however, we simply want to be able to port a parallel loop to GPU without writing lots of boilerplate code. We also want the code to be able to work with different accelerators. *Directive-based approaches* can help with this: original code is annotated with *directives*, **`#pragma`**'s in C/C++, which indicate the intent executing a fragment of code on GPU, and it is up to the compiler to take care of details. The most widely used set of directives for accelerators is OpenACC [13]; OpenMP 4.0 [14] now also includes accelerator support, but as of 2013, this is supported not by all compilers. OpenACC directives are supported by PGI, HMPP and Cray C/C++ and Fortran compilers. A simple program for matrix-vector multiplication using OpenACC is presented in Fig. 13; note that it is significantly shorter than the kernel introduced previously. `parallel` directives denotes the parallel region, while `copyin` and `copyout` directives indicate which, where to and how much data to copy. This is a very simple OpenACC example, where the compiler can automatically detect loop paralellism; in more complex cases, the developer has to use

`independent` directive to indicate that a loop is parallel. More advanced OpenACC features include data regions, which control lifetime of data, asynchronous execution, specifications of loops with independent iterations and interoperability with functions written in CUDA.

```
void matvecmul_gpu
(float * restrict b, const float * restrict a,
const float * restrict x, int n) {
  #pragma acc parallel copyin(a[0:n*n], x[0:n]) copyout(b[0:n])
  for(int i = 0; i < n; i++) {
    b[i] = 0;
    for(int j = 0; j < n; j++)
      b[i] += a[j * n + i] * x[j];
  }
}  // matvecmul_gpu
```

Fig. 13: *Code for matrix-vector multiplication using OpenACC*

The plot in Fig. 14 gives some idea about performance of matrix multiplication using approaches described previously. The comparison is done on Kepler K40X GPU with CUDA 5.5. OpenACC, CUDA and cuBLAS reach 44, 293 and 1222 GFlop/s (double precision), respectively, with peak device performance being 1430 GFlop/s. In all cases, only time spent in doing operations on GPU is counted, not the time spent in copying data between host and device. cuBLAS variant, specifically tuned for GPU hardware, obviously performs the best, achieving 85% of peak performance. This demonstrates the reason to use specifically tuned GPU libraries when they are available: it allows to get significant performance without writing any GPU code. Simple CUDA implementation using shared memory is much slower, but it achieves 20% of peak performance. The CUDA version can be improved, by e.g. computing several elements in a single GPU thread; after heavy tuning, it is possible to reach performance levels of cuBLAS. OpenACC version is currently the slowest; but it can still be improved by fine-tuning annotation parameters.

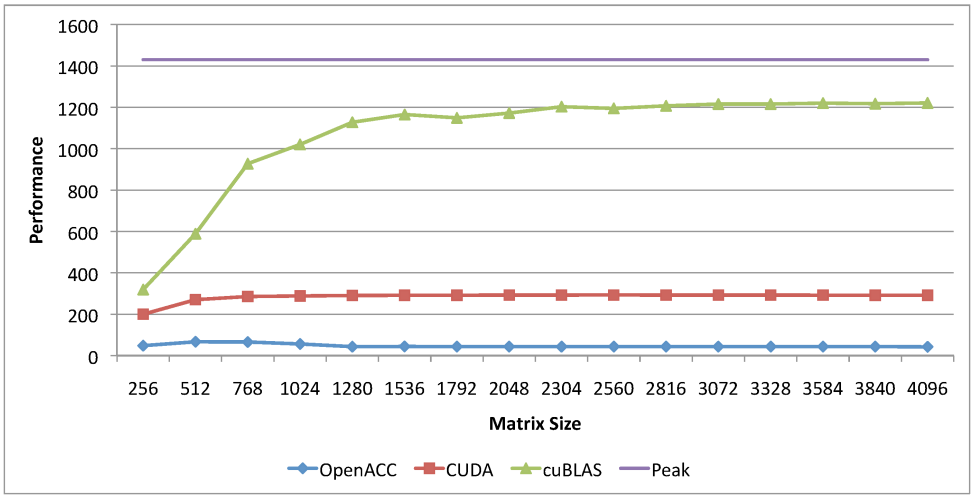


Fig. 14: *Comparison of cuBLAS, CUDA and OpenACC versions of matrix multiplication*

4.5 Debugging and Performance Analysis

The ecosystem of NVidia GPUs now includes tools for debugging and performance analysis. Two of them come directly with CUDA Toolkit: cuda-gdb and NVidia Visual Profiler. **cuda-gdb** is a debugger based on GDB providing for debugging programs running on GPU. Its abilities include setting breakpoints inside GPU kernels, and viewing the state of execution of different GPU threads currently running. **NVidia Visual Profiler** is a profiler, which by default collects timings of CUDA API calls and CUDA kernels. It can compose the list of kernels taking most of the time, print a trace of CUDA API calls or present a timeline view of kernel launches and CUDA API calls with multiple devices and streams. In addition, NVidia Visual Profiler can collect detailed metrics, such as number of memory transactions or percentage of cache hits, to help identify problems in the application. **NVidia NSight** is an integrated development environment (IDE) for GPU development, available as plug-ins both for Microsoft Visual Studio and Eclipse. NVidia NSight provides integration for CUDA syntax highlighting, GPU debugging and profiling, which together aims at making GPU development more productive.
NVidia also provides **CUDA Profiling Tools Interface (CUPTI)**, which enables third-party applications to collect metrics similar to NVidia profiler. Tools like **Vampir** [15] and **Score-P** [16] use this API to collect data, and then combine it with data from other sources, e.g. MPI profiling, to build their own traces. Such tools can be extremely useful when identifying performance problems in applications running on GPU clusters.

5 Conclusion

In this lecture, we discussed GPU programming. We first presented a brief history of GPUs, and discussed their architecture. We then described CUDA, the main API currently used to program GPUs, and presented an example of matrix-vector multiplication on GPU using CUDA. We then discussed other approaches to programming GPU, as well as performance analysis and debugging tools for GPU and libraries and applications with GPU support.
As the volume of these lecture notes is rather limited, we were able to only briefly touch GPU programming, and describe the most important aspects. If you would like to do GPU programming yourself and are searching for a good book, we would recommend "Programming Massively Parallel Processors: a Hands-on Approach" [18]. It starts from GPU architecture and programming basics, but covers a wide range of advanced topics, which are illustrated through samples and sometimes through real applications. It is relatively up to date and covers even most recent GPU features. Another good starting point is NVidia Documentation website [17], which contains CUDA C Programming GUIDE and also comprehensive documentation of CUDA API and various accompanying libraries. It is also a good idea to download CUDA Toolkit and look through the samples for examples of GPU programming. More advanced reading includes GPU Computing Gems series [19], [20], which contains lots of examples of GPU techniques and applications from various fields. Talks on specific applications can also at [21], and [22] is a good place for tips and tricks.
GPU programming is a very young and fast evolving field. Just six years ago, when CUDA was first released, GPU supported only a single kernel at a time and only single-precision computation. Current GPUs support double precision, can access host memory, run multiple kernels simultaneously with copying data, and, with the latest devices, even launch new kernels from GPU. Future evolution will likely include closer integration of GPUs with the host CPU and sharing the same memory between both. Companies such as NVidia, Intel and AMD are work-

ing in that direction. It is hard to say now how a future accelerator would look like, but one point is clear: many-core accelerators have become an integral part of our computing environment, whether discrete or part of the CPU. This equally applies to mobile and embedded systems, PCs and supercomputers. And it is necessary to know how to make use of them in order to achieve maximum performance on future computers.

References

[1] H. Meuer, E. Strohmaier, J. Dongarra, H. Simon. Top 500 Supercomputer Sites. `http://www.top500.org` (2013).

[2] E. Elsen, M. Houston, V. Vishal, E. Darve, P. Hanrahan, V. Pande. N-Body Simulations on GPUs. Proceedings of 2006 ACM/IEEE Conference on Supercomputing.

[3] M. Segal, K. Akeley. The OpenGL Graphics System: A Specification, Version 2.0 (2004).

[4] NVidia Corporation. CUDA C Programming Guide, Version 5.5 (2013).

[5] NVidia Corporation. GPU Applications. `http://www.nvidia.com/object/gpu-applications.html` (2013).

[6] A. Munshi. The OpenCL Specification, Version 1.2 (2012).

[7] Altera Corporation. Altera SDK for OpenCL Programming Guide (2013).

[8] AMD Corporation. Heterogeneous Computing Webpage. URL: `http://developer.amd.com/tools-and-sdks/heterogeneous-computing/` (2013).

[9] Matrix Algebra on GPU and Multicore Architectures (MAGMA). URL: `http://icl.utk.edu/magma/` (2013).

[10] Open Source Computer Vision Library (OpenCV). URL: `http://opencv.org` (2013).

[11] M. A. Clark, R. Babich, K. Barros, R. Brower, C. Rebbi. Solving Lattice QCD Systems of Equations using Mixed Precision Solvers on GPUs (2010).

[12] J. Kraus. An Introduction to CUDA-Aware MPI. NVidia Parallel Forall blog. URL: `http://devblogs.nvidia.com/parallelforall/introduction-cuda-aware-mpi/` (2013).

[13] OpenACC Corporation. The OpenACC Application Programming Interface, Version 2.0 (2013).

[14] OpenMP Architecture Review Board. OpenMP Application Program Interface, Version 4.0 (2013).

[15] Vampir — Performance Optimization: Vampir Website. URL: `http://www.vampir.eu` (2013).

[16] Score-P — Scalable Performance Measurement Infrastructure for Parallel Codes Webpage. URL: http://www.vi-hps.org/projects/score-p/ (2013).

[17] NVidia Corporation. NVidia CUDA Documentation Website. URL: http://docs.nvidia.com/cuda (2013).

[18] D. Kirk, W. W. Hwu. Programming Massively Parallel Processors: A Hands-on Approach (Morgan Kaufmann, 2012).

[19] W. W. Hwu. GPU Computing Gems Emerald Edition (Applications of GPU Computing Series) (Morgan Kaufmann, 2011).

[20] W. W. Hwu. GPU Computing Gems Jade Edition (Appilcations of GPU Computing Series) (Morgan Kaufmann, 2011).

[21] NVidia Corporation. GTC On-Demand Webpage. URL: http://www.gputechconf.com/gtcnew/on-demand-gtc.php (2013).

[22] NVidia Corporation. Parallel ForAll Blog.
URL: http://devblogs.nvidia.com/parallelforall/ (2013).

D 6 DFT on Supercomputers [1]

R. Zeller
Institute for Advanced Simulation
Forschungszentrum Jülich GmbH

Contents

1 Introduction

Since the early days of quantum mechanics it has been recognized that the quantum mechanical equations, except for the simplest systems, can be solved only with the help of advanced numerical techniques and computational devices. This even applies if the full quantum mechanical problem is reduced to the much simpler task of determining the electronic density necessary in density-functional theory. Thus practically from its beginning with the Thomas-Fermi approximation in 1927, the development of density-functional methods always has been connected deeply with the development of computing devices. The achievements of this co-development have been honoured in a pronounced manner by the Nobel Prize in Chemistry 1998 which was divided equally between Walter Kohn "for his development of the density-functional theory" and John A. Pople "for his development of computational methods in quantum chemistry".
The aim of the present lecture is to provide an account of the permanent connection between supercomputing and density-functional theory starting with a historical perspective, explaining the current situation and ending with a consideration of some important issues for future calculations. The main topics to be covered are two recent algorithmic developments which are capable of using hundreds of thousands of parallel processors as they are available on supercomputers today. These topics concern calculations in real space, where the locality of the physical systems is directly exploited in the local communication pattern available on supercomputers, and calculations with linear-scaling complexity, where some accuracy is traded in order to overcome the usual bottleneck of density-functional calculations which consists in the fact that the computing time increases proportionally to the third of the number of atoms in the system. These topics will be exemplified with two recent codes (ju*RS* and KKR*nano*). These codes were produced during the Ph.D. thesis works of Paul Baumeister [1] and Alexander Thieß [2] at the Peter Grünberg Institute and the Institute for Advanced Simulation at the Forschungszentrum Jülich with the design aim that they should be applicable in an efficient manner for massively parallel computations on present and future supercomputers.

2 Historical Development

It is fair to start this section with Paul Dirac's widely cited introductory sentences from his paper "Quantum Mechanics of Many-Electron Systems" [3] which read "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble. It therefore becomes desirable that approximate practical methods of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation."
Since that time, Dirac's paper appeared in 1929, an enormous number of approximations and methods have been invented and used by mathematicians, physicists, chemists, engineers and others to obtain this goal often with the discovery that the possible work was limited by the computing devices available. It is interesting to see that this apparently was recognized already by Eugene Wigner and Frederik Seitz when they developed the cellular method for bandstructure calculations of periodic solids. In their 1933 paper "On the Constitution of Metallic Sodium" they point out that "The calculation of a wave function took about two afternoons, and five wave functions were calculated on the whole, ..." This remark might be understood as a comment that

such heavy calculations could be done for simple systems like alkali metals, but were too time consuming for the treatment of more complicated systems.
For about two decades until the 1950s the progress in electronic structure calculations remained slow mainly perhaps because of the lack of computing devices that were better than humans equipped with mechanical or electromechanical desktop calculating machines. So even in 1952 and 1953, when Frank Herman [4, 5] predicted by calculations with the method of orthogonalized plane waves (OPW) that diamond and germanium had multi-valley conduction band structures with indirect bandgaps, these calculations were extraordinarily time consuming as he recollects in an article published in Physics Today [6]. There he writes "I showed my mother how to set up the OPW secular equations and how to factor them, and she agreed to do some of this in her spare time in order to save me time. ... So she did the factoring by hand–the total effort may have taken her many hundreds of hours–and I was saved many weeks, if not months of detailed work."
The situation changed dramatically when the first electronic "supercomputers" appeared like the ENIAC which was in operation from 1946 to 1955. Suddenly the computing power was increased by factor 1000. Although the ENIAC (an abbreviation which stands for "Electronic Numerical Integrator And Computer") had been developed for military purposes like the calculation of ballistic trajectories and for numerical weather prediction, the great potential of such machines was clearly recognized by solid state physicists at that time. They saw that computing with electronic devices offered a unique opportunity to test theoretical concepts by sufficiently accurate numerical calculations. Thus they acted also as a driving force for the advancement of electronic computing devices as John Slater recollects in his 1967 article "The Current State in Solid-State and Molecular Theory" [7] on pages 75 to 83. In this article he also summarizes: "It was the development of these machines in the post-war period, first slowly, then at an ever increasing rate, which has had perhaps the most striking effect on the direction of the quantum theory of molecules and solids since the war".
After the 1950s with the phenomenal advances in numerical electronic structure calculations, the progress was not so rapid in the 1960s and at the beginning of the 1970s, on the one hand due to doubts of how exchange and correlation effects could be incorporated accurately into one-electron calculations and on the other hand due to the prejudice expressed by many physicists that computational work does not belong to physics but to engineering. Here the situation changed dramatically at the end of the 1970s, in particular by the work of Jepsen and Andersen [8], Moruzzi, Janak and Williams [9] and Yin and Cohen [10]. This work clearly showed that even the simple local density approximation for exchange and correlation could be used to obtain nearly quantitative agreement with experiment. Another stimulation came from the discovery that even in metallic systems the concept of a one-electron bandstructure was not only a theoretical construct, but could be confirmed by high-resolution angle-resolved photoemission spectroscopy [11] in good agreement with earlier theoretical predictions [12].
Since then an ever increasing amount of density-functional calculations has been published. This shows up, for instance, in the almost 3000 citations that the two density-functional papers of Hohenberg and Kohn of 1964 and Kohn and Sham of 1965 together have obtained in the year 2013 alone, and in the fact that first-principles simulations within density-functional theory have become a major consumer of supercomputing resources in many computing centres around the world. It is remarkable that this work is not only impressive because of its quantitative amount, but also impressive because of the highly efficient codes which have been developed over the years and which have been honoured by numerous Gordon Bell Prizes, which recognize outstanding achievement in high-performance computing each year. The first Gordon Bell Prize

for density-functional calculations was awarded in 1990 for work entitled "Electronic structure of a high-temperature superconductor" [13] which was done for the perovskite superconductor $Ba_{1-x}K_xBiO_3$ on an Intel iPSC/860 using 128 parallel processors and the up to now last one in 2011 for work entitled "First-principles calculations of electron states of a silicon nanowire with 100,000 atoms on the K computer" [14], which was done using 663552 parallel processors.

3 Present Situation

For several decades the growth in computing power was mainly due to an increase of the processor speed on single processor machines. In the 1980s further growth in computing power was realized by multiprocessor shared-memory machines as the Cray X-MP which had two processors with a combined peak performance of 400 MFLOPS. Another step forward was the use of "massive parallel" distributed-memory computers as the Intel Paragon and the Cray T3E with up to about 2000 processor elements in the 1990s, a trend which continues until today where the largest supercomputers are equipped with hundreds of thousands of processors. Today the most powerful supercomputer of the world, the Tianhe-2, is equipped with 3.12 million processors. A further development is the use of multi-processor computing nodes with a common shared memory for the processors on each node. An alternative is the incorporation of accelerator units into a compute node.

Compared to single-processor or shared-memory multiprocessor machines, distributed-memory machines require explicit consideration of the data communicated between the processors. Too much communication seriously impairs the speedup which can be gained by using more and more processors. Obviously, up to now this challenge was mastered very well by the density-functional community as can be seen from the above mentioned Gordon Bell Prizes.

For the present and the future, the ever increasing number of processors seems to require considerable attention. To utilize hundreds of thousands of processing elements, it is essential to use an approach that has good parallel efficiency. This goal seems to be difficult to achieve by traditional plane-wave codes because of large communication requirements for systems with thousands of atoms, the present target of density-functional calculation on supercomputers. Therefore real-space codes, which will be discussed in section 4, are becoming more and more the choice for such large calculations.

Another issue, which must be considered seriously for density-functional calculations for large systems, is the increasing heterogeneity of supercomputers. The traditional single-level parallelization approach (as implemented via the MPI communication library) must be replaced very likely by a multi-level strategy where MPI parallelization is supplemented by multi-threaded OpenMP directives and/or special programming techniques for attached accelerator units.

Finally, there is also as fundamental issue which must be overcome if density-functional calculations are intended for larger and larger systems. Traditionally, the main computional effort consists in the solution of algebraic eigenvalue problems with matrices of a dimension which increases linearly with the number N of atoms in the system. This leads to computing times which increase as N^3. This increase represents a serious bottleneck even if petascale computing eventually is replaced by exascale computing. With a factor of 1000 in computing power only ten times more atoms can be treated which means that the diameter of the systems which can be treated can increase only by a factor $\sqrt[3]{10} \approx 2$. One way to overcome this N^3 bottleneck could be the use of so-called linear-scaling techniques which will be discussed in section 5.

4 Real-Space Calculations

In order to deal with the communication bottleneck, it is necessary to reduce communication time as much as possible. In this respect, traditional plane-wave calculations for periodic solids, which are based on fast Fourier transform (FFT) techniques, are not the most effective approach. FFT needs all-to-all communications and the communication time grows with increasing number of processors. Therefore, the present trend goes into the direction of grid-based real-space algorithms that exploit the intrinsic locality of the real-space Kohn-Sham equations. The values of wave functions, electron density, and potential field at the grid points are stored locally and the amount of inter-processor communication is kept low by exchanging data only between nearest-neighbour processors which are connected on the three or more dimensional torus network of the supercomputer.
The basic parallelization strategy in grid-based real-space methods is domain decomposition over the real-space grid. In addition trivial parallelization can be applied over the two spin directions in collinear magnetic systems and over the k-points necessary for the Brillouin zone sampling if the considered systems are represented by periodic supercells. Since for very large systems usually only one k-point is used, another possible level of parallelization is over the electronic states (bands). The band parallelization, however, is not trivial because the matrices, which must be diagonalized to obtain the bands from the eigenstates, contain information which must be collected from all grid points and not only from the ones available on the local processor.
Using this strategy, in recent years several grid-based real-space density-functional codes have been developed for supercomputers, for instance, in Japan [14], Finland [15] and Germany [1]. As an example, the code ju*RS*, which was developed during the Ph.D. thesis of Paul Baumeister [1] at the Forschungszentrum Jülich, will be discussed in some detail in following subsection.

4.1 Example: ju*RS*

The real-space finite-difference code ju*RS* is based on the projector augmented wave method [16] which models the correct scattering of the Kohn-Sham wave functions at the atomic cores and provides accurate all-electron energy contributions of the atom. ju*RS* was designed to run on massively parallelized computers with the aim to do structural optimization of systems with more than 1000 atoms. It treats densities, potentials and Kohn-Sham wave functions on uniform three-dimensional grids with a straightforward parallelization using domain decomposition of the grid points where each domain is assigned to one compute node. ju*RS* also supports highly efficient spin and k-point parallelization.
Because the number of grid points, which determine the size of Hamiltonian matrix, is much larger than the number of the lowest (occupied) eigenstates necessary to determine the density and because the Hamiltonian matrix is sparse due to the locality of the discretized Kohn-Sham equations, the use of a direct eigensolver is prohibitively more expensive than the use of an iterative one. For the solution, ju*RS* applies the direct inversion of iterative subspace (DIIS) method [17], which can be done for each state independently, together with the subspace rotation (SR) technique in the space of occupied bands in order to avoid time-consuming explicit Gram-Schmidt orthogonalization.
If it is necessary because of memory limitations or if it is desired because of faster execution time, the parallelization over spin, k-points and domains can be supplemented with a a third level of parallelism by distributing the bands into band sets and assigning each band set to

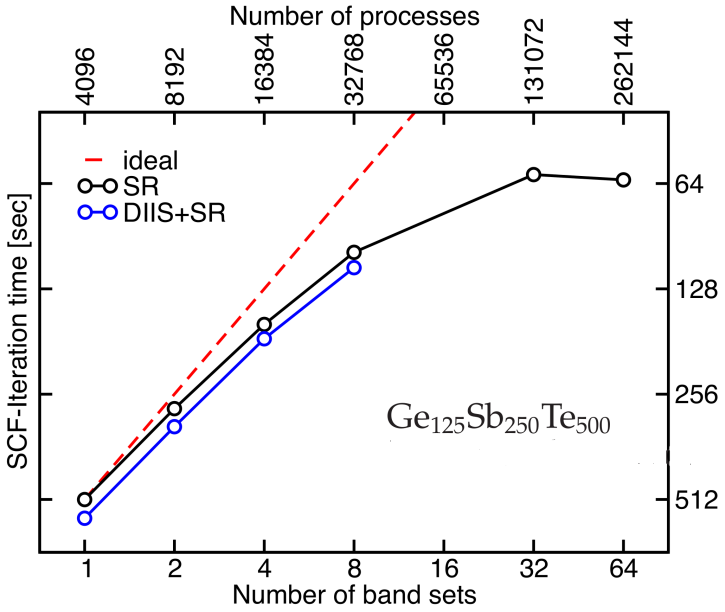


Fig. 1: *Strong scaling in the parallelization over bands for a system of 875 atoms. With increasing number of processors the execution time on JUGENE decreases for up to a 32-fold parallelization. Note that the time axis is inverted so that the curves give an indication of the obtained speedup. (Picture taken from Ref. [1], p. 102.)*

a different processor. Because of the distributed storage and treatment of the bands this is a strongly communicating task which certainly limits the degree of efficient parallelization which can be achieved. Nevertheless, band parallelizaion can be used with success as shown in Fig. 1, where the execution (wall-clock) time for one selfconsistency step is plotted for a system of 875 atoms as function of the number of processors used for each domain. The underlying grid parallelization for the considered system was over $16 \times 16 \times 16 = 4096$ domains so that in total 4096 to 262144 processors were used for the results shown in the figure. With increasing number of processors the execution time decreases up to the use of 32 band sets while no further decrease was achieved with 64 band sets. Note that 262144 processors represent 89 % of all precessors available on the Blue Gene/P computer JUGENE which was in operation from 2007 to 2012 at the Jülich Supercomputing Centre. Note also that the time axis is inverted in the figure and that the broken red curve indicates the ideal speedup which would be obtained in case of 100 % parallel efficiency.

The computer code juRS has been applied so far, for instance, for the investigation of how the electronic structure changes in $GeSb_2Te_4$ during the phase transition from a crystalline to an amorphous state. $GeSb_2Te_4$ belongs to the so-called phase-change materials which are of importance for technological applications as rewritable DVD discs or non-volatile random access memory because of the fact that the transition from the crystalline to the amorphous state can be switched back and forth by laser pulses or electric currents and can be detected by optical or resistivity measurements.

The crystalline state of $GeSb_2Te_4$ is characterized by the rock-salt crystal structure where the sites on one sublattice are occupied by Te atoms. The sites on the other sublattice are occupied by Ge and Sb atoms more or less randomly and one quarter of these sites remains empty. The vacancies in $GeSb_2Te_4$ have a profound influence on the resistivity properties. Of particular importance is the arrangement of the vacancies around the Te atoms because the local density of

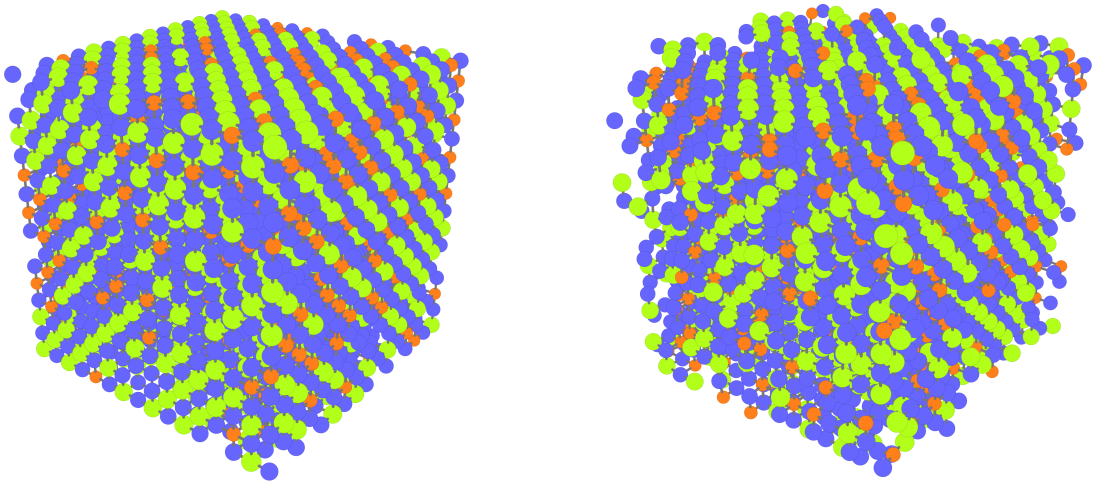


Fig. 2: *Visualization of the atoms in the random alloy $Ge_{512}Sb_{1024}Te_{2048}$ in the ideal initial lattice positions (left) and after relaxation (right). Blue, green and orange balls indicate Te, Sb and Ge atoms. (Pictures are taken from Ref. [1], p. 128 and p. 130.)*

states (LDOS) at the Fermi level increases prominently with an increasing number of vacancy neighbours so that these vacancy clusters act as scattering centres for electronic transport. A visualization of the atomic positions in a 3584-atom supercell is given in Fig. 2 where on the left side atoms are shown in the initial ideal lattice positions and on the right side in the final positions after structural optimization was performed with the juRS code. An important consequence of the structural relaxation for the electronic structure of $GeSb_2Te_4$ consists in a reduction of the LDOS at the Fermi level with an associated reduction of the strength of the scattering centres.

5 Linear-Scaling Calculations

In order to deal with the N^3 bottleneck, it is necessary to find and implement algorithms with less than $O(N^3)$ computational complexity. One way to proceed is to avoid the direct solution of the algebraic eigenvalue problem and to replace it by an iterative solution. If the Hamilitonian matrix is made sparse, which is possible by using localized basis sets or real-space grids, and if the orthogonalization of the eigenstates is avoided, which is possible by different ways, as for instance described in Ref. [18], the computational complexity is reduced to $O(N^2)$. Another way to proceed is to use the so-called nearsightedness principle of electronic matter [19]. This concept means that in systems without long range electric fields the electronic density at a point in space does not depend in any relevant way on potential changes far away. This concept directly leads to the so-called linear-scaling methods with complexity $O(N)$ independent of whether iterative or direct solution techniques are used.

A simple application of the nearsightedness concept consists in dividing the entire system into subsystems, calculating the density in each isolated subsystem and combining these densities to obtain the density of the entire system approximately. This approach is used, for instance, in the original divide-and-conquer method [20], in the locally self-consistent multiple-scattering (LSMS) method [21] for which Gordon Bell Prizes were awarded in 1998 and 2009, in the charge-patching method (CPM) and in the linearly scaling three dimensional fragment (LS3DF)

method [22] for which a Gordon Bell Prize was awarded in 2008.
While the above kind of methods can be build on existing codes and methods, there exist also linear-scaling methods based on entirely new implementations using sparse Hamiltonian and iterative solution techniques together with the nearsightness principle, for example the codes CONQUEST and ONETEP which were developed in Great Britain [18] and which can be used for accurate density-functional calculations for tens or hundreds of thousands of atoms.
At present it is, however, still difficult for newly developed linear-scaling methods to compete with the widely used plane-wave density-functional method which have been developed intensively over the at least last three decades. But for very large systems linear-scaling methods seem to be the only ones which can be successful on future supercomputers with many million of processors because they can be implemented efficiently for massively parallel computing. As an example, the code KKR*nano*, which was developed during the Ph.D. thesis of Alexander Thieß [2] at the Forschungszentrum Jülich, will be discussed in some detail in following subsection.

5.1 Example: KKR*nano*

Contrary to most other codes, KKR*nano* does not solve the differential Kohn-Sham wavefunction equation, but an integral equation, the so-called Dyson equation, where the integral kernel consists of a product of the Kohn-Sham effective potential and the Green function a suitably chosen reference system. The solution of the Dyson equation directly gives the Kohn-Sham Green function from which the fundamental quantity of density-functional theory, the electronic density, is obtained by contour integration in the complex energy plane.
The main computational effort in KKR*nano* consists in the standard linear algebra problem of solving large systems of linear equations. These equations are obtained from the integral equation by dividing the integral, which extends over the space covered by the system, into integrals over space-filling, non-overlapping cells around the atoms and by using the analytically known angular momentum representation of the free space Green function in each cell.
The difference of KKR*nano* compared to conventional programs based on the Korringa-Kohn-Rostoker (KKR) multiple-scattering method is that the systems of linear equations are not solved directly, for instance by Gauss elimination with an effort proportional to N^3, but by the iterative quasi-minimal residual (QMR) method of Freund and Nachtigal [23]. This is advantageous for massively parallel computing because the QMR iterations can be done independently for each atom and each angular moment channel and because only sparse matrix multiplications are required if the properties of the repulsive reference system [24] of the screened KKR method are exploited. Compared to dense matrices with $O(N^2)$ elements the sparse matrices in the screened KKR method have only $O(N)$ non-zero elements. Because of that the QMR iterations for each atom can be done with $O(N)$ floating point operations so that the total effort increases only as N^2. Note that this reduction from N^3 is not connected with any loss of accuracy [25] if the residual error of the QMR iterations is made small enough.
The natural parallelization in KKR*nano* is domain decomposition with a simple matching between atomic cells and processors. Communication is necessary only at the beginning of a self-consistency step, when single-cell t matrices (usually with only 16×16 matrix elements), which are calculated independently from potential in each cell, must be exchanged, and at the end of the self-consistency step, when charge multipole moments, which are calculated independently from the density in each cell, must be exchanged to obtain the electrostatic potential needed in the next self-consistency step. In addition to the parallelization over the atomic cells,

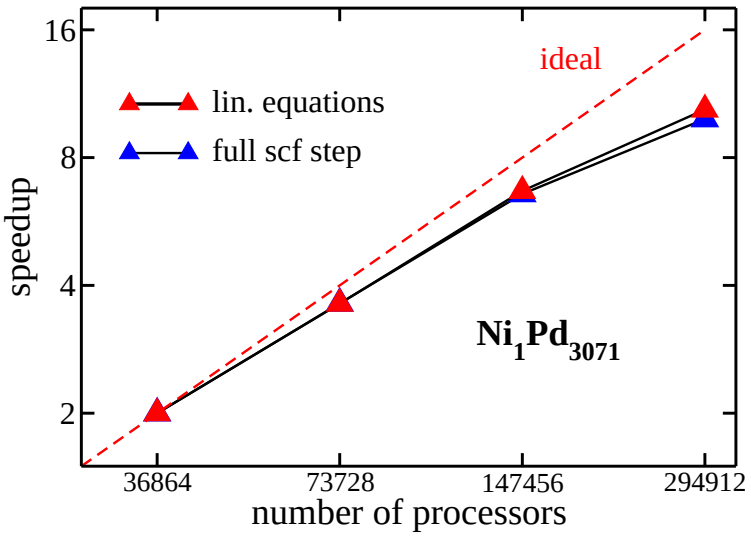


Fig. 3: *Strong scaling in the parallelization over the angular momentum channels for a system of 3072 atoms. With increasing number of processors a speedup is obtained up to all 294912 processors available on JUGENE.*

KKR*nano* can be used with three other levels of MPI parallelization. These parallelization are over groups of integration mesh points on the complex energy contour, over the two spin directions in magnetic systems and over the (usually 16) angular momentum channels. Together, the four levels of MPI parallelization make KKR*nano* suitable for hundreds of thousands of processors as it is shown in Fig. 3. Here for a NiPd system of 3072 atoms the strong-scaling speedup is shown which was gained if the solution of the linear equations is parallelized using the innermost parallelization over the angular momentum channels. The curve with the red symbols is for the time required for the solution of systems of linear equations, while the curve with the blue symbols is for a full-selfconsistency step. The very similar behaviour of the curves shows that the overhead from the (not yet over the angular momentum channels parallelized) rest of the program is small. Even with increasing the number of processors from 147456 (using half the number of processors available on the Blue Gene/P computer JUGENE at the Jülich Supercomputing Centre) to 294912 (using all processors on JUGENE) a speedup (of about 1.5) was obtained. Thus it is not unreasonable to assume that KKR*nano* can utilize millions of processors if systems with much more atoms than in Fig. 3 are studied.

The computer code KKR*nano* has been applied so far, for instance, to study effects of vacancy clustering on the superparamagnetism in Gd-doped GaN [26], to study effects of substitutional Ni and Cu defect atoms on the martensitic phase transition in the shape memory alloy NiTi [27] and to study the role of vacancies in metal-insulator transitions of the phase-change material $GeSb_2Te_4$ [28]. Here one aim was to investigate whether the interpretation of the metal-insulator transition in terms of a disorder induced Anderson localization, which was proposed in order to explain the experimental results, could be confirmed by density-functional calculations. A finite-size scaling study was performed using supercells with up to 8000 sites and, although these supercells were still somewhat too small, the interpretation in terms of Anderson localization could be supported. The other aim was to obtain insight into the observed

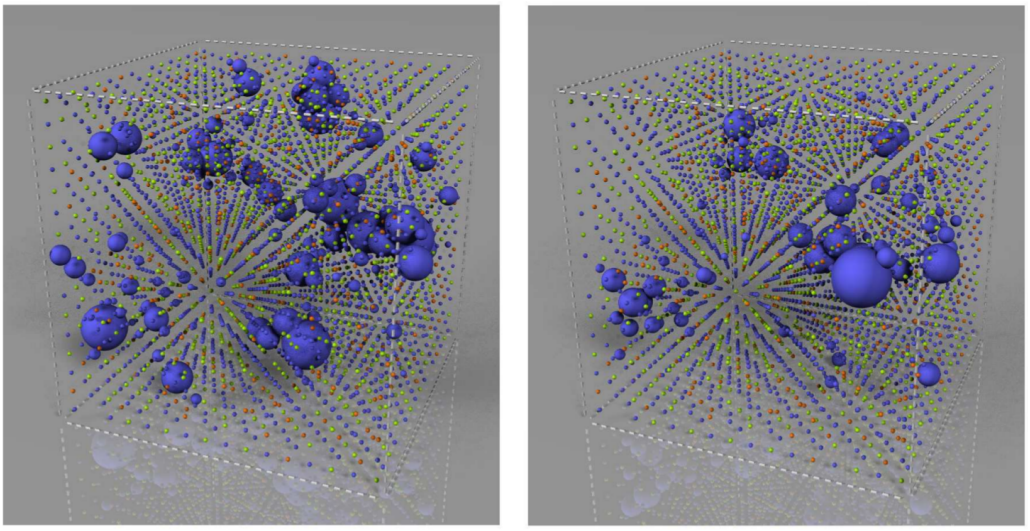


Fig. 4: *Spatial distribution of the local density of states (LDOS) at the Fermi level for a* $Ge_{512}Sb_{1024}Te_{2048}$ *supercell with 512 vacancies. Size ond colour of the balls indicate height of the LDOS and kind of the atoms, blue for Te, green for Sb and orange for Ge. The left picture is for the initial random configuration and the right picture is for the configuration obtained after simulated annealing. (Figure is taken from Ref. [2], p. 158.)*

pronounced dependence of the metal-insulator transition on the annealing temperature. Here simulated annealing by a kinetic Monte-Carlo technique was performed where the configuration probabilities were determined from the local energy values which can be calculated in the KKR method from energy density within each atomic cell. Fig. 4 shows how the LDOS at the Fermi level changes between the initial random configuration of the atoms and the configuration obtained for a temperature of 1000 K after 5×10^8 Monte-Carlo time-steps. The main effect of the annealing, which becomes more effective with higher temperatures, is that vacancy clusters with more than four vacancies are dissolved. This leads to less regions with high LDOS and thus to a reduction of the number and strength of the scattering centres.

A very important feature of KKR*nano* for future calculations of very large systems is that the concept of nearsightedness can easily be exploited by choosing a truncation region around each atom and neglecting matrix elements of the Kohn-Sham Green function outside of the truncation region. Then the computing time for the QMR iterations is independent of system size and the same is true for the memory necessary for each atom, i.e. for each processor. This makes calculations with hundreds of thousands of atoms feasible, for instance, on the Blue Gene/Q at the Jülich Supercomputing Centre. For this purpose, at present the code in KKR*nano* is reorganized by Elias Rabel as part of his Ph.D. work. A typical example for the behaviour of the execution time in linear-scaling mode is shown in Fig. 5. Without truncation the time increases almost linearly which indicates that the quadratically scaling total computational work is well distributed over the increasing number of processors. With truncation the execution time is much smaller and increases only slightly with system size indicating that the overall computational effort practically increases only linearly with increasing N. This means that with increasing number of processors in the future an increasing number of atoms can be treated without the need to wait longer and longer until the results are calculated.

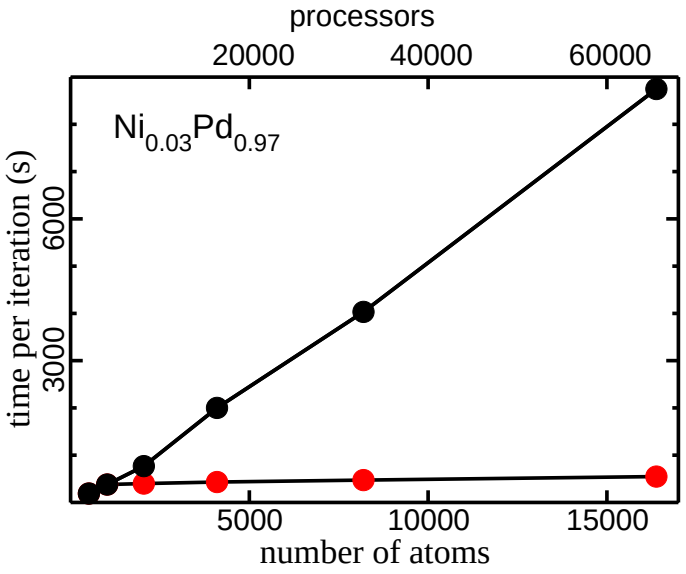


Fig. 5: *Execution time on JUGENE for one self-consistency step for a dilute NiPd alloy (with a concentration of 3 % randomly placed Ni atoms) as function of system size without truncation of the Green function (upper curve with black dots) and in linear-scaling mode for a truncation region of 959 atoms (lower curve with red dots). Four processors per atom were used.*

6 Ongoing Challenges

The future of density-functional theory on supercomputers will be affected certainly by the hardware design of future supercomputers, the algorithmic development of density-functional codes for these supercomputers and the availability of sufficient developer teams for this task.
The problem with hardware design is apparent from the expected increase of the number of processor into the many million range and from the increased complexity due to many-processor chips with internal single-instruction-multiple-data (SIMD) capabilities or due to attached accelerator units. This requires a change from the present typical single-parallelization approach focused on pure MPI realization to a multilevel approach using multithreaded code with efficient exploitation of low level SIMD instructions and to the use of special software for the accelerator units. A particularly important issue in hardware design is the amount of input and output data which can be handled in a reasonable timeframe. Although density-functional calculations are rather modest in this respect because essentially only input densities must be read before the self-consistency steps and only output densities must be stored after these steps, there are rather large intermediate wave-function data sets which could be used for check-pointing in case of hardware faults. Fortunately, both problems have been realized by the manufacturers of future supercomputer and can be dealt with by the paradigm of co-design where hardware and application software are developed in consideration of each other.
The problem with algorithmic and code development for new computing architectures arises from the long development time of simulation packages, which can reach into tens to hundreds of man-years, and the long lifetime of these packages, which typically extend over more than ten years and sometimes over decades. This is certainly longer than the lifetime of previous and probably also future supercomputer architectures and requires not only that the software packages should be highly portable to changing architectures, but also that architectures should

not change so rapidly that work invested in these packages becomes obsolete. Fortunately, this problem has been realized also by the manufacturers of future supercomputer and can be dealt with by the same paradigm of co-design.

The problem for scientists and teams, which develop and improve density-functional codes, arises from the fact that novel algorithm and codes have development times of about five to ten years, which exceeds the length of many funding initiatives, and that students and postdocs committed to development work cannot publish as many papers as others who simply apply existing codes to systems of highly topical interest. Moreover, papers on development work are often not suitable for high-impact journals, for instance in Nature and Science, because scientific breakthroughs due to software tools often appear only years after the tools have been developed, adopted and validated. This problem has existed in the past, but will become probably considerably more important in the future due to the increasing hardware complexity. The appreciation of this fact by funding agencies and persons responsible for the careers of young scientists would promote certainly the opportunities which arise from future exascale computing capabilities.

Despite the above mentioned problems, however, it can expected that density-functional calculations on supercomputers will be an important tool for supporting materials design and development for advanced 21st century applications. This expectation can be given with rather large confidence because the past has shown that people working with density-functional theory have acquired always remarkable expertise in writing highly efficient codes and because this trend very likely will continue so that simulations of complex systems consisting of several thousand atoms or more will become routinely possible in the not so distant future.

References

[1] P. F. Baumeister, *Real-Space Finite-Difference PAW Method for Large-Scale Applications on Massively Parallel Computers*, Ph.D. thesis, RWTH Aachen University (2012).

[2] A. R. Thieß, *Development and application of a massively parallel KKR Green function method for large scale systems*, Ph.D. thesis, RWTH Aachen University (2011).

[3] P. A. M. Dirac, Proc. R. Soc. Lond. A **123**, 714 (1929).

[4] F. Herman, Phys. Rev. **88**, 1210 (1952).

[5] F. Herman and J. Callway, Phys. Rev. **89**, 518 (1953).

[6] F. Herman, Phys. Today **37**, 56 (1984).

[7] J. C. Slater, Int. J. Quant. Chem. **1**, 37 (1967).

[8] O. Jepsen, O. K. Andersen, and A. R. Mackintosh, Phys. Rev. B **12**, 3084 (1975).

[9] V. L. Moruzzi, A. R. Williams, and J. F. Janak, Phys. Rev. B **15**, 2854 (1977).

[10] M. T. Yin and M. L. Cohen, Phys. Rev. Lett. **45**, 1004 (1980).

[11] P. Thiry, D. Chandesris, J. Lecante, C. Guillot, H. Pinchaux, and Y. Pétroff, Phys. Rev. Lett. **43**, 82 (1979).

[12] G. A. Burdick, Phys. Rev. **129**, 138 (1963).

[13] G. A. Geist, B. Ginatempo, W. A. Shelton, and G. M. Stocks, J. Supercomput. **6**, 153 (1992).

[14] Y. Hasegawa, J.-I. Iwata, M. Tsuji, D. Takahashi, A. Oshiyama, K. Minami, T. Boku, H. Inoue, Y. Kitazawa, I. Miyoshi, and M. Yokokawa, Int. J. High Perf. Comput. Appl. (2013), published online first, doi:10.1177/1094342013508163.

[15] J. Enkovaara, C. Rostgaard, J. J. Mortensen, J. Chen, M. Dulak, L. Ferrighi, J. Gavnholt, C. Glinsvad, V. Haikola, H. A. Hansen, H. H. Kristoffersen, M. Kuisma, *et al.*, J. Phys.: Condens. Matter **22**, 253202 (2010).

[16] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[17] D. M. Wood and A. Zunger, J. Phys. A **18**, 1343 (1985).

[18] D. R. Bowler and T. Miyazaki, Rep. Prog. Phys. **75**, 036503 (2012).

[19] E. Prodan and W. Kohn, Proc. Natl. Acad. Sci. USA **102**, 11635 (2005).

[20] W. Yang, Phys. Rev. Lett. **66**, 1438 (1991).

[21] Y. Wang, G. M. Stocks, W. A. Shelton, D. M. Nicholson, Z. Szotek, and W. M. Temmerman, Phys. Rev. Lett. **75**, 2867 (1995).

[22] L.-W. Wang, Annu. Rev. Phys. Chem. **61**, 19 (2010).

[23] R. W. Freund and N. M. Nachtigal, Num. Math. **60**, 315 (1991).

[24] R. Zeller, P. H. Dederichs, B. Újfalussy, L. Szunyogh, and P. Weinberger, Phys. Rev. B **52**, 8807 (1995).

[25] R. Zeller, J. Phys.: Condens. Matter **20**, 294215 (2008).

[26] A. Thiess, P. H. Dederichs, R. Zeller, S. Blügel, and W. R. L. Lambrecht, Phys. Rev. B **86**, 180401(R) (2012).

[27] R. Zeller, Phys. stat. solidi **xx**, xxxxxx (2014).

[28] W. Zhang, A. Thiess, P. Zalden, R. Zeller, P. H. Dederichs, J.-Y. Raty, M. Wuttig, S. Blügel, and R. Mazzarello, Nature Mater. **11**, 952 (2012).

Index

The IFF Spring School & Computational Materials Physics in Jülich

PGI-JCNS-TA
52425 Jülich
Phone: *++49 2461 61-1739*
Fax: *++49 2461 61-2410*
E-Mail: *springschool@fz-juelich.de*
web: *www.iff-springschool.de*

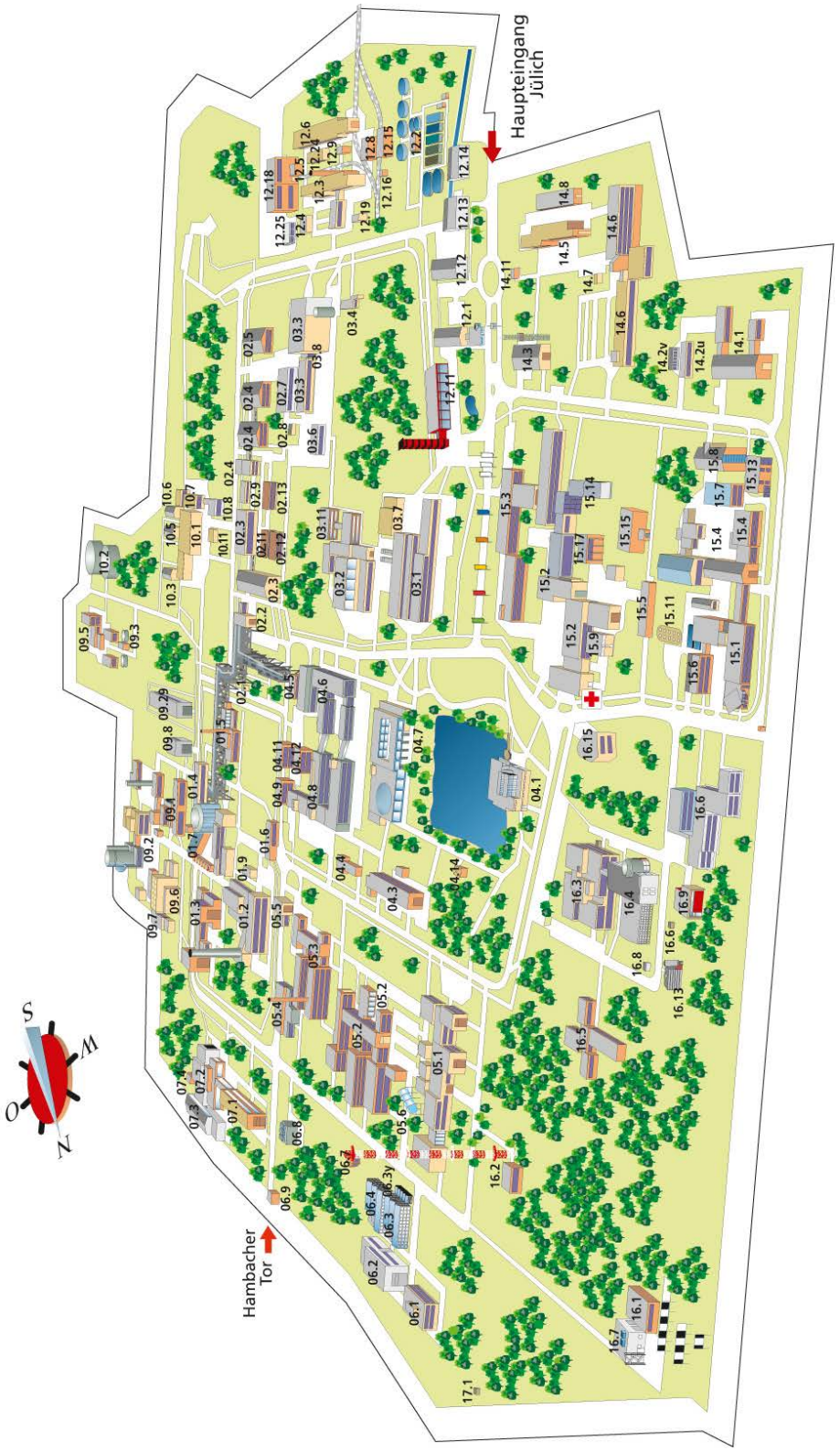
The IFF Spring Schools were first brought into being in 1970 by the "Institut für Festkörperforschung" (IFF). Since then, they have made it possible for students and young scientists to gain a two-week insight into a current topic related to condensed matter physics. In 2011, IFF was dissolved as part of a restructuringprocess within Forschungszentrum Jülich, and new institutes, namely the "Peter Grünberg Institut" (PGI), the "Jülich Centre for Neutron Science" (JCNS) and the "Institute of Complex Systems" (ICS) were established. Together with the "Institute for Advanced Simulation" (IAS), they will continue to coordinate the IFF Spring Schools. The 45th Spring School 2014 was organized by PGI-1/IAS-1 "Quantum Theory of Materials" together with the Theoretical Solid State Physics Group from the RWTH Aachen University.

Forschungszentrum Jülich has a long track record of advancing research in the field of computational materials physics and materials science. At a very early stage, scientists here recognized the importance of computer simulation as a source of knowledge in condensed matter physics. With the advent of peta-flop computing based on new architectural concepts, scientific computing is transforming to simulation sciences. Together with the Jülich Supercomputing Centre (JSC), the German Research School for Simulation Sciences (GRS), established under the umbrella of the Jülich-Aachen Research Alliance (JARA), and the recently founded Institute for Advanced Simulation (IAS), Forschungszentrum Jülich is home to a research triangle where education, different research disciplines together with methodical and technological competences are combined to manage the future challenges in simulation sciences.

The Institute "Quantum Theory of Materials" develops conceptual and computational methods based on density-functional theory. It supports an "ab initio" simulation laboratory. Computational materials science research is established by combining first-principles results with dynamical and statistical methods (molecular dynamics, Monte Carlo). The institute focuses on the analysis and computation of electronic properties of solid-state systems relevant for nanoelectronics (spintronics, nanoionics, complex magnetism, organic molecules) and energy.

The development of quantum many-body methods and their application in model systems of correlated electrons has a long history at the physics department of RWTH Aachen University. Nowadays, the department is split into the institutes of Theory of Statistical Physics and Theoretical Solid State Physics. The research pursued in the former is focused on correlation effects in mesoscopic systems including the rapidly developing field of non-equilibrium phenomena, using renormalization group-based approaches. In the latter, such methods as well as numerical tools (e.g. quantum Monte Carlo) are used to investigate the physics of correlated materials.

Forschungszentrum Jülich - Campus Map



Band / Volume 61
Study of intermolecular interactions in hetero-organic thin films
B. Stadtmüller (2013), viii, 198 pp
ISBN: 978-3-89336-871-6

Band / Volume 62
Structure, magnetism and excitations in some Mn-based magnetocaloric effect compounds
M. Gottschlich (2013), 175 pp
ISBN: 978-3-89336-874-7

Band / Volume 63
Neutron Scattering
Lectures of the JCNS Laboratory Course held at Forschungszentrum Jülich and the research reactor FRM II of TU Munich
edited by Th. Brückel, G. Heger, D. Richter, G. Roth and R. Zorn (2013),
ca 350 pages
ISBN: 978-3-89336-880-8

Band / Volume 64
Neutron Scattering
Experiment Manuals of the JCNS Laborator Course held at Forschungszentrum Jülich and the research reactorFRM II of TU Munich
edited by Th. Brückel, G. Heger, D. Richter, G. Roth and R. Zorn (2013),
ca. 150 pages
ISBN: 978-3-89336-881-5

Band / Volume 65
Solutions of Exercises in "An Introduction to Dynamics of Colloids"
J. K. G. Dhont, K. Kang (2013), 143 pp
ISBN: 978-3-89336-882-2

Band / Volume 66
Strukturelle und chemische Charakterisierung von selbst-assemblierten Monolagen organischer Moleküle auf Oberflächen
M. Müller (2013), 168 pp
ISBN: 978-3-89336-883-9

Band / Volume 67
Adsorption of (hetero-)organic Phthalocyanine and PTCDA thin films on Ag(111)
C. Kleimann (2013), iv, 149 pp
ISBN: 978-3-89336-888-4

Band / Volume 68
High-energy high-momentum surface spin waves of ultrathin epitaxial 3d transition metal films
R. Jayaraman (2013), ix, 135 pp
ISBN: 978-3-89336-890-7

Band / Volume 69
Magnetic Oxide Heterostructures: EuO on Cubic Oxides and on Silicon
C. Caspers (2013), xiii, 153 pp
ISBN: 978-3-89336-891-4

Band / Volume 70
Femtosecond Spin Dynamics in Magnetic Multilayers Employing High Harmonics of Laser Radiation
D. Rudolf (2013), vi, 121 pp
ISBN: 978-3-89336-894-5

Band / Volume 71
Development and application of a massively parallel KKR Green function method for large scale systems
A. R. Thieß (2013), ii, 173 pp
ISBN: 978-3-89336-906-5

Band / Volume 72
Conformational Dynamics of Calmodulin and Ribosome-Nascent Chain Complexes Studied by Time-Resolved Fluorescence Anisotropy
P. Lamprou (2013), 182 pp
ISBN: 978-3-89336-907-2

Band / Volume 73
Stereoselective synthesis of vicinal diols with enzymatic cascade reactions
J. Kulig (2013), XV, 177 pp
ISBN: 978-3-89336-911-9

Band / Volume 74
Computing Solids
Models, ab-initio methods and supercomputing
Lecture Notes of the 45th IFF Spring School 2014
March 10 – 21, 2014 Jülich, Germany
ed. by S. Blügel, N. Helbig, V. Meden, D. Wortmann (2014), ca. 1000 pp
ISBN: 978-3-89336-912-6